FINAL TECHNICAL REPORT

OFFICE OF NAVAL RESEARCH

GRANT #N00014-86-K-0401 1 July 1986 - 30 June 1989
GRANT #N00014-89-J-3055 1 July 1989 - 31 December 1995

PRINCIPLES OF PERCEPTION IN BAT SONAR

James A. Simmons
Department of Neuroscience
Box 1953
Brown University
Providence, RI 02912

1994

TEL: (401) 863-1542 FAX: (401) 863-1074 email: james_simmons@brown.edu

1. Introduction

This report describes the scientific findings of a program of research supported by the Office of Naval Reseach on echo processing and acoustic imaging by the biological sonar of bats. Results and conclusions cover the performance of bat sonar in critical benchmark tasks, the auditory neurophysiological mechanisms for echo reception and signal processing, and the computational basis for transforming waveforms of sonar broadcasts and echoes into acoustic images. The research identifies the nature of the algorithms bats use to process echoes and suggests methods for emulating critical aspects of performance in man-made systems.

1.1 Bats and Echolocation

Bats are nocturnal flying mammals classified in the order Chiroptera. These animals have evolved a biological sonar, called *echolocation*, to orient in darkness--to guide their flight around obstacles and to detect their prey (Griffin 1958; Neuweiler 1990; Novick 1977; see Popper and Fay 1995). Echolocating bats broadcast ultrasonic sonar signals that travel outward into the environment, reflect or scatter off objects, and return to the bat's ears as echoes. First the outgoing sonar signal and then the echoes impinge on the ears to act as stimuli, and the bat's auditory system processes the information carried by these sounds to reconstruct images of targets (Dear et al. 1993; Dear, Simmons, and Fritz 1993; Schnitzler and Henson 1980; Simmons 1989; Simmons and Kick 1984; Suga 1988, 1990).

1.2 The Big Brown Bat

The big brown bat, *Eptesicus fuscus* ("dusky house-flier"), is a common, widely-distributed North American bat of the family Vespertilionidae (Kurta and Baker 1990). *Eptesicus* is one of many species of insectivorous bats that produce *frequency-modulated* (FM) echolocation sounds and use echoes to find and intercept flying insects (Griffin 1958; Neuweiler 1990; Pye 1980; Simmons 1989; see Popper and Fay 1995). Figure 1 shows an insect's-eye view of a big brown bat as it approaches a target during an interception maneuver guided by sonar. A wide range of behavioral experiments have been carried out with *Eptesicus* to evaluate aspects of the performance of its sonar (see Moss and Schnitzler 1995 and Simmons et al. 1995). These studies identify fundamental features of the auditory computations underlying FM echolocation by specifying the final output of these computations--the *images*--in relation to the input--*emissions and echoes* (Simmons 1989, 1992; Saillant et al. 1993).

1.3 Scope of this Report

This report describes auditory computations which create the dimension of distance, or *target range*, in the images perceived by the big brown bat and the physiological mechanisms which support these computations. To provide a conceptual framework, we use a model of echolocation (**S**pectrogram **C**orrelation **a**nd **T**ransformation--SCAT; see Saillant et al. 1993) which assumes that the bat's cochlea (1) segments the range of frequencies in the bat's sonar sounds into parallel band-pass-filtered channels, (2) half-wave-rectifies and then smooths the resulting frequency segments of the sounds, and (3) triggers neural discharges from these excitation patterns. The bat's sonar sounds are frequency-modulated (FM), and the model uses a modified *auditory spectrogram* format (see Altes 1980, 1984) for initial encoding of their frequency sweeps, with several parameters (*e.g.*, scaling of filter center frequencies, sharpness of tuning, integration-time) quantified from physiological data. Taking "neural discharges" triggered from the auditory spectrograms of broadcast and received sounds as input to further computations, the model focuses attention on reconstruction of the locations of echo-sources

along the axis of target range to form images equivalent to the "A-scope" display of a radar or sonar system (essentially a plot of target range *versus* target strength; see Skolnik 1962).

2. *The Sonar System of the Big Brown Bat*

2.1 *Echolocation Sounds used by* Eptesicus

Bats universally seem to use FM echolocation signals as acoustic probes to determine target range from the travel-time of echoes (Griffin 1958; Schnitzler and Henson 1980; Simmons 1973; Simmons and Grinnell 1988). *Eptesicus* broadcasts ultrasonic sounds that contain frequencies in the range from about 20 kHz to over 100 kHz depending on the bat's situation. Figure 2 shows spectrograms of echolocation signals recorded during the bat's approach to a small target in an interception maneuver (Simmons et al. chapter in Popper and Fay in press). These sounds are harmonically-structured FM signals, with first and second harmonics always present (FM_1, FM_2; terminology after Suga 1988, 1990), plus a segment of the third harmonic (FM_3) and often also a segment of the fourth harmonic (FM_4) present, too. Usually, FM_1 sweeps from about 50-60 kHz down to about 20-25 kHz, while FM_2 sweeps from about 100 kHz down to 40-50 kHz (Fig. 2a-d). Only a short segment of FM_3 is produced around 75-90 kHz, and FM_4 is also restricted to around 80-90 kHz when it appears (Fig. 3e-g). In Figure 2, the shapes of the FM sweeps are curved, with the sweep-shape being approximately *hyperbolic*, making the sweep itself approximately *linear with period*. That is, as *frequency* sweeps curvilinearly downward from 55 kHz to 23 kHz in FM_1, *period* sweeps upward from about 18 μs to 43 μs in a linear fashion (see below).

2.2 *Target Ranging*

The bat's sonar sounds travel outward from the bat's mouth to impinge on objects located at different *distances* and return to the bat's ears as echoes at different *times*. The bat perceives the distance to objects from this time delay (Simmons 1973). In air, the delay of echoes is 5.8 ms per meter of range--the time required for the sound to travel over the two-way path-length of *2 m* for a target at a range of *1 m*. Figure 3 illustrates the relation between the target's location at a particular range (*r*) and the duration of the bat's sonar sounds. As a rule-of-thumb, *Eptesicus* keeps the duration of its sounds slightly shorter than the two-way travel-time of echoes (echo delay, *t*, in Fig. 3) by progressively shortening the sounds during interception, so there is little or no overlap of transmissions with echoes from the insect (Griffin 1958; Hartley 1992; Simmons 1989; see Simmons et al. chapter in Popper and Fay in press). The distance in the air spanned by the sound's duration almost completely fills up the path-length from the bat's mouth out to the target and back to its ears (Fig. 3).

2.3 *Operating Range of Echolocation*

Eptesicus can detect insect-sized targets as far away as 5 m, which corresponds to an echo delay of about 30 ms (Kick 1982). Small objects located farther away than this maximum operating range return echoes that are too weak to be heard (Lawrence and Simmons 1982; Pye 1980). At long range, *Eptesicus* can detect echoes at levels as low as 0 dB SPL (Kick 1982; Kick and Simmons 1984). The bat's broadcasts are roughly 100-110 dB stronger than the weakest echo that can be detected, so the bat can tolerate considerable attenuation of echoes due to the small size or long range of targets before the echoes become inaudible. Consequently, a wide range of different sizes of insects or other objects located at different distances potentially are detectable with echolocation, which means that the bat's target-ranging system must be able to accommodate reception of echoes at a variety of different delays and display objects a a variety of different distances. The axis of echo-delay (target range) in the "A-scope" images perceived by *Eptesicus* must extend from roughly 0.5 ms (about 10 cm) to roughly 30 ms (about 5m) (Dear, Simmons, and Fritz 1993; Simmons 1989).

3. *Targets, Echoes, and Images*

3.1 Target Glints and the Structure of Echoes

Objects in air behave as though they consist of a small number of parts that each reflect a more-or-less faithful replica of the incident sonar sound (Simmons and Chen 1989). Consequently, the overall "echo" that actually returns to the bat's ears from a natural target actually consists of several discrete echoes arriving together. Figure 4 shows a two-glint target (a dipole) as the simplest example of a complex object. The reflecting surfaces or points in the target are called *glints* (*A* and *B* in Fig. 4), and their separation from each other in range □ḍr) determines the time separation (ḍt) of the reflected replicas. The flying insects that bats prey upon are small; different body-parts (head, abdomen, wing-tips, legs) are separated from each other by only 2-3 cm, so the delay separations between their reflections will be less than 100-150 μs. The critical factor is the relatively small size of these delay separations in relation to the duration of the incident sonar sound. Because the distance from the bat to the insect is large compared to the separation of the insect's parts, and because the bat uses this overall distance to determine the duration of its sonar sounds (Fig. 3), the multiple reflections sent back to the bat will overlap each other when they arrive (Fig. 4).

3.2 "Auditory Spectrograms" of Emissions and Echoes

3.2.1 Frequency Scale and Integration-Time

Figure 2 shows spectrograms of echolocation sounds that have a conventional vertical axis that is linear with frequency. The FM sweeps of the bat's sonar sounds appear curved on this scale because the sweep functions are approximately hyperbolic, or linear with period. However, the bat's auditory system does not scale frequency as a linear variable. Physiological measurements reveal that frequency scaling in the auditory system of *Eptesicus* is approximately hyperbolic, too (see below). Figure 5 shows an example of an "auditory" spectrogram with a hyperbolic frequency axis for a sonar sound of *Eptesicus* (*emission*) and two echoes (*A* and *B*) arriving at different delays (t_A = 3.7 ms and t_B = 4.7 ms). The duration of the broadcast signal is 2 ms, and it has two harmonics (FM_1,FM_2; see Fig. 2). The shape of each sweep in the emission and both echoes is "linearized" in the auditory spectrogram by the hyperbolic vertical frequency scale. These spectrograms serve as the initial signal representations in the Spectrogram Correlation and Transformation (SCAT) model of echolocation (Saillant, et al. 1993). They are made by passing the signals through a bank of 81 parallel band-pass filters with constant, 4-kHz bandwidths and center frequencies spaced at fixed period intervals (*1/f*) of 0.5 μs. The filter outputs are half-wave rectified and then smoothed with a low-pass filter of about 1 kHz to approximate at least some of the physiological events, associated with transduction, which produce neural discharges from excitation comparable to the horizontal slices in Figure 5.

The two echoes in Figure 5 are only 1 ms apart, while the broadcast sound is 2 ms long, so the raw waveforms of the echoes overlap each other (top of Fig. 5). However, the *auditory spectrograms* of the echoes do *not* overlap because the time-width or *integration-time* of the spectrogram (time-width of spectrogram slices in Fig. 5) is only a fraction of a millisecond (about ±300 μs to ±400 μs). The two echoes in Figure 5 will always appear as separate spectrograms with discrete FM sweeps as along as their time-separation (ḍt) exceeds the integration-time (Altes 1984; Beuter 1980).

3.2.2 Auditory Spectrograms of Overlapping Echoes

Figure 6 illustrates hyperbolically-scaled auditory (*SCAT*) spectrograms for a series of overlapping echoes (*A* and *B*) at different time separations (ḍt) from 0 to 1 ms. These echo-delay separations correspond to separations of 0 to about 17 cm along the axis of range. The shorter separations of 0, 60, and 110 μs are realistic values for the echoes reflected by different parts of a small target, such as an insect, with dimensions of up to about 2 cm, while larger time separations of 225 μs to 1 ms correspond better to the range separations associated with different targets, such as an insect located anywhere from 4 cm to 17 cm away from background vegetation. Figure 6 demonstrates the

importance of the integration-time associated with the initial stages of auditory processing for determining how information about targets must be represented in the bat's auditory system (Altes 1984; Beuter 1980; Simmons 1989). For time-separations of 350 μs and greater, the two overlapping echoes are represented by separate SCAT spectrograms, with a recognizable sweep for each harmonic (FM_1 and FM_2 in Fig. 5). However, for separations of less than 350 μs, the separate spectrograms merge into only one recognizable sweep for each harmonic (Simmons et al. 1989). However, now the *amplitude* of the spectrogram is modulated at different frequencies by interference between the echoes as they mix together within the integration-time of the parallel filter channels. That is, the spectrum of the combined echoes is rippled by the interference. At longer echo separations of 450 μs and more, the spectrograms in Figure 6 appear as smooth, sloping ridges, but at shorter separations of 60, 110, and 225 μs, the ridges contain peaks and valleys--spectral features--whose separation in *frequency* (vertical axis) provides the only indication that the echoes actually are separated in *time* (horizontal axis). The integration-time of spectrograms thus establishes a boundary between representation of the time separation of the two echoes as a difference *in time* along the horizontal axis of the spectrograms, and representation of the time separation as a difference *in frequency* between peaks and valleys along the vertical axis (Simmons 1992; see Moss and Schnitzler chapter and Simmons et al. chapter in Popper and Fay in press).

3.2.3 The Bat's Integration-Time for Echo Reception

In clutter-interference experiments, where the bat's task is to detect test echoes at one particular arrival-time in the presence of masking echoes at different arrival-times, *Eptesicus* fails to detect the test echoes when the time separation is smaller than 350 μs (Simmons et al. 1989). This result reveals that the bat receives and segregates echoes with an intrinsic integration-time of about 350 μs. Echoes that arrive closer together than this time window become merged into a single sound for purposes of detection. Physiological responses recorded in the bat's auditory system also fail to register overlapping sounds at short time separations. Neurons in the cochlear nucleus and nucleus of the lateral lemnisicus (see Fig. 9) can register the presence of two separate sounds with separate discharges as long as these sounds are over 300-500 μs apart (Covey and Casseday 1991; Grinnell 1963; Suga 1964); for shorter time separations the neurons fail to produce discharges in response to the second sound by itself. It thus appears that *Eptesicus* fails to detect test echoes in the presence of cluttering echoes because its auditory system merges the two sounds together and represents them both with just one volley of neural discharges.

3.3 Segregation of Echoes within the Integration-Time

Because insects are small objects, the reflected replicas from their glints will arrive closer together than the integration-time of 350 μs and will overlap to interfere with each other (Kober and Schnitzler 1990; Moss and Zagaeski 1994; Schmidt 1992; Simmons and Chen 1989). Do bats perceive the actual range separation of the glints in the insect (Simmons 1989, 1992, 1993; Simmons, Moss, and Ferragamo 1990) or just the spectral effects of interference between the overlapping reflections (Neuweiler 1990; Schmidt 1992)? If the two glints in Figure 4 are just two different parts of the same insect, it might seem unnecessary for the bat to perceive that there really are two glints at slightly different ranges in order to intercept the target. The time delay associated with the spectrogram of the combined reflections in the echo from the insect would be adequate to perceive the insect's overall distance, and the spectrum of the overlapping echo replicas ought to be adequate to characterize the target's shape and fluttering motion without explicitly perceiving the range separation of its glints (Neuweiler 1990; Schmidt 1992; Simmons and Chen 1989).

However, the same configuration of closely-spaced echoes often occurs in situations where perception of the range to each glint probably *is* necessary. It is perfectly possible for the two echo

replicas in Figure 4 to come from different objects rather than from the same object. For example, the nearer glint (*A*) might be part of an insect while the farther glint (*B*) might be part of some vegetation that the insect happens to be flying past. In this situation the bat would need to perceive the distance to the insect (r_A) in order to affect a capture, while it would also need to perceive the distance to the vegetation (r_B) to avoid colliding with it. Just characterizing the overlapping echoes by the interference spectrum would not be enough to accomplish these two tasks; the bat has to perceive *both* ranges to catch the insect while avoiding the obstacle (Simmons et al. chapter in Popper and Fay in press). The problem is that information about the ranges of the two glints has been blurred by the integration-time of 350 μs so that *both* echoes have only *one* delay which can be determined directly from the time axis of the spectrogram (Fig. 6).

3.5 The Bat's Images of Two-Glint Targets

A crucial aspect of performance in echolocation must be the bat's ability to resolve two closely-spaced echoes as arriving at different times. What is the smallest separation in the arrival-time of two echoes (ḍt in Fig. 4) that the bat can perceive? Analysis of the behavior of bats in several naturalistic situations (obstacle-avoidance tests, interception of targets in clutter, discrimination of airborne or suspended targets) reveals that even a flying bat probably can resolve two echoes as having discrete delays for separations as short as 5-20 μs (see Simmons et al. chapter in Popper and Fay in press). Figure 7 illustrates the results of experiments which demonstrate more directly that *Eptesicus* can perceive the arrival-times of both replicas contained in two-glint echoes at small delay separations. The graphs show the performance (% errors) of two bats detecting changes in the arrival-time of overlapping pairs of test echoes separated by 0, 10, 20, or 30 μs (*dt*) at an overall delay of about 3.2 ms. In this experiment, probe echoes are moved to different arrival-times in relation to either of the test echoes (located at 0 μs and at 10, 20, or 30 μs on horizontal axis of graphs in Fig. 7); the decline in the bat's performance when the probe echoes are aligned in time with either of the test echoes (peaks in error curves) shows that the bat characterizes these double echoes as having two integral delay values, not just one delay value. (For details of experiments and compound performance plots, see Simmons et al. 1990.) In effect, Figure 7 shows examples of the *"A-scope" images* perceived by *Eptesicus* for two-glint targets with range separations □ḍr) of 0, 1.7, 3.4, or 5.2 mm.

The most important feature of the images in Figure 7 is that the two overlapping echoes are both assigned perceived magnitudes of delay along a scale of delay subdivided into sufficiently fine steps that differences of 10, 20, or 30 μs can be displayed. The smallest echo-delay separation that *Eptesicus* can perceive with a separate error peak in the image for each delay (as in Fig. 7) is about 2 μs (Saillant et al. 1993). The scale of delay in the bat's images must therefore be graduated in increments no larger than 2 μs, otherwise the two echoes would have been assigned the same delay value. *Eptesicus* is extraordinarily accurate at perceiving small changes in the arrival-time of echoes: Several experiments demonstrate that the bat can detect changes smaller than 0.4-0.5 μs (Menne et al. 1989; Moss and Schnitzler 1989; Simmons 1979), and the smallest detectable change actually measured is about 10-15 ns (Simmons, et al. 1990). It is thus plausible that the spacing of adjacent "units" of delay along the echo-delay axis in the bat's images really could be as small as 1-2 μs.

4. Computations on Spectrograms to Determine Delay and Recover Resolution

4.1 The Computational Problem in Echolocation

4.1.1 Fine Delay Resolution, but Coarse Integration-Time

The challenge for understanding the auditory computations which support echolocation lies in the difficulty of achieving fine temporal resolution of 10, 20, or 30 μs for multiple echoes arriving within the 350-μs integration-time of echo reception (Fig. 7). These computations have to produce an

accurate estimate of the arrival-time of the first reflected replica in each echo while also overcoming the limitations imposed by the *blurring* effects of the integration-time to produce an estimate of the arrival-time of the second reflected replica, too (Saillant et al. 1993; Simmons 1989; Simmons et al. 1990). Because echoes separated by 10, 20, or 30 μs are merged into only one spectrogram (Fig. 6), an estimate of delay for only one of the echoes can be made directly from the spectrograms. The estimate of delay for the other echo has to be derived instead from the effects of overlap and interference on the merged spectrograms. Crucially, the *inputs* to the computations which yield the two delay estimates in each image in Figure 7 must have different numerical formats--one essentially a measurement of time between the emission and echo spectrograms, and the other essentially a measurement of frequencies for spectral peaks and notches. However, *both* estimates of delay are manifested along the same dimension of the images (Fig. 7), so the *outputs* of the underlying computations ultimately must converge upon the same numerical scale in the bat's images. Moreover, the graduations of the echo-delay axis in these images must be finely-divided enough to allow echo-delay resolution down to about 2 μs (Saillant et al. 1993). Our goal is to learn how the bat's images are created, and identification of the computational locus for this convergence of formats would place us close to the site of image-formation itself.

4.1.2 Convolution *to Form Spectrograms*

During band-pass filtering by the inner ear, the FM waveform is *convolved* with the impulse-response of each filter, and the output, which consists of a segment of the waveform of the FM sweep in the neighborhood of the filter's center frequency, is then half-wave rectified and smoothed by a low-pass filter with a cut-off frequency of about 1 kHz to produce discharges of auditory-nerve fibers. The smoothing filter is the most significant limiting component in this peripheral auditory signal-conditioning regime; for most practical purposes, the time-of-occurrence of each frequency in the FM sweep comes to be represented by the impulse-response of this smoothing filter, which is about 300-400 μs wide. Subsequently, in the nervous system, these peripheral, receptor-generated impulses are themselves replaced by the neural discharges they trigger; these discharges also are spike-shaped pulses several hundred microseconds wide. Each horizontal slice of the auditory spectrogram for the emission or the echo is approximately the width of this impulse-response (see *integration-time* in Fig. 5). When two echoes arrive close enough to each other that these impulse-responses collide and form a single impulse, the volleys of neural discharges they produce also merge into one volley, and the bat can no longer detect one echo in the presence of the other. The clutter-interference experiment measures the time-window for the collision of these impulses as viewed by the bat, giving a value of about 350 μs as the minimum separation required for the bat to detect one echo as a separate sound in the presence of the other.

4.1.3 Deconvolution *to Form Images*

The result of convolution is to replace the series of frequency-time points in each FM sweep with a sloping ridge about 300-400 μs wide (Fig. 5). To segregate echoes whose sweeps overlap and merge into just one set of ridges in the spectrogram (Fig. 6), the bat has to *deblurr* these sweeps-as-ridges to recover the temporal resolution that was lost during convolution. This deblurring operation is called *deconvolution*, and it requires explicit knowledge of the transmitted waveform so that the presence of multiple replicas can be recognized even when they arrive so close together that they merge into one spectrogram (Fig. 6). The key to understanding the computations at the heart of echolocation lies in knowing what is meant by this "knowledge" and how it can be used to reverse the blurring effects of convolution.

4.2 The SCAT Model

Figure 8 is a diagram of the principal computational stages required to convert the raw time-series waveform of a sonar emission and two overlapping echoes into an "A-scope" sonar image

depicting the delay of both echoes along the same scale of time. This diagram shows the signal-processing operations of the SCAT model as a guide to identifying what has to be learned about echo-processing in the bat. The model's first stage is its "cochlea" (Saillant et al. 1993), which uses 81 band-pass filters in parallel to transform the raw input waveforms into hyperbolically-scaled SCAT spectrograms (see Figs. 5-6) for further processing. This component of the model emulates the most critical features of the bat's inner ear and then generates "neural discharges" registering successive frequencies in the FM sweeps of the sounds. The model's remaining stages are two parallel pathways--a *spectrogram correlation* system for determining the time separation between the spectrogram of the emission and the spectrogram of echoes (Altes 1980, 1984), and a *spectrogram transformation* system for converting the pattern of peaks and notches in the spectrogram of overlapping echoes (Beuter 1980; Altes 1984) into an estimate of the time-separation of the merged replicas. The outputs of these two processing pathways converge to write values of echo delay along a single delay axis in the final images.

4.3 Delay-Lines for Spectrogram Correlation to Determine Echo Delay

4.3.1 Storing the Shape of the FM Sweeps in Emissions for Comparison with FM Sweeps in Echoes

Figure 8A shows the raw waveforms of the input signals--a sonar transmission with a duration of 2 ms and two overlapping echoes (*A* and *B*). The delay of the first echo (t_A) is 3.7 ms, and the delay of the second echo (t_B) is only 60 μs larger. This short delay separation ($\dot{d}t$ = 60 μs) results in the two echoes merging to form just one spectrogram in Figure 8B. The amplitude of the echo spectrogram at different frequencies contains peaks and notches reflecting the interference that takes place within the 350-μs integration-time of the spectrograms. As a first step, the SCAT model determines the arrival-time of the compound echo (*A* + *B*) from the time-intervals between the spectrogram of the emission and the spectrogram of the echo at different frequencies (that is, the horizontal time displacement of the echo spectrogram to the right of the emission spectrogram in Fig. 8B). These *spectrogram delays* (shown as t_{A1}-t_{A5} in Fig. 8B) are extracted using *delay-lines* that register the time-of-occurrence of each frequency in the emission and then compare it with the time-of-occurrence of the corresponding frequency in the echo. In effect, the delay-lines store the shape of the spectrogram for the emission and slide it to the right in Figure 8B until it lines up with the spectrogram for the echo--a process equivalent to correlation of the echo and emission spectrograms (Altes 1980). An "event" travels along each delay-line from one delay tap to the next to register the occurrence of one specific frequency in the broadcast sweep, and the relative position of similar events across all the delay-lines preserves the shape of the sweep as the events propagate along the delay-lines. This property of the spectrogram is crucial because the sweep-shape really represents information about the *phases* of the different frequencies in the broadcast sound. If the shape of the sweep in the echo is distorted in the course of reflection from the target (the frequencies in the echo undergo different phase-shifts), and this change in shape is detected, then whatever target feature caused the change in shape can be incorporated into the range image.

4.3.2 Reading Echo Delay from Delay Taps

At each frequency, the delay of the echo (t_{A1}-t_{A5}) is represented by the specific delay tap in the delay-line that is active at the same moment that the echo arrives. This "moment" is judged by detecting coincidences between events taking place at the delay taps and events triggered by the incoming echo. The spectrogram delays (t_{A1}-t_{A5}), which are represented by the active delay taps in different delay-lines, are then averaged across all the delay-lines in Figure 8C to estimate the delay of the echo as a whole (t_A). This overall delay value is obtained from measurements of the timing of discharges and represents the distance to the object that contains the two glints; it usually is interpreted to be the distance to the nearer of the two glints (Saillant et al. 1993; Simmons et al. 1990; Simmons 1993). In the absence of noise, all 81 channels normally register their delay estimates at the same delay value (or

delay-line tap); the addition of noise merely broadens the distribution of active delay taps around the mean value, and, at high levels, noise sometimes displaces the mean, too. Registration of echo delay is very precise by this method when averaged across a number of parallel delay-lines, and *Eptesicus* is also very accurate at determining the delay of echoes to within 10-15 ns from the timing of neural discharges (Simmons et al. 1990). Delay-lines and delay-coincidence devices are commonly used in radar systems to display the arrival-times of echoes. Furthermore, this part of the model is equivalent to the delay-coincidence-correlation process used in some models of auditory pitch coding (Langner 1992; Licklider 1951), and a delay-coincidence model specifically tailored to echolocation is widely assumed to be the basis for perception of target range in bats (Park and Pollak 1993; Sullivan 1982; Suga 1988, 1990; see below).

If the processing of echo information were to stop at this point, the range image would depict just the overall distance to the target. No distinction would be made about the distances to the two glints. Further information about the target is contained in the shape of the spectrum of the overlapping echoes (Fig. 8D), and one widely-accepted hypothesis is that the bat classifies targets in terms of the spectral coloration supplied by the peaks and notches at different frequencies (Neuweiler 1990; Schmidt 1992). The bat does not appear to stop at this point, however, because it perceives both delays associated with the two-glint target in the same image (Fig. 7).

4.4 Spectrogram Transformation to Determine Delay Separation

4.4.1 Knowledge of Signals for Deconvolution

The capacity to deconvolve two overlapping echoes that have been merged into one spectrogram (Fig. 8A-B) depends upon being able to translate the pattern of peaks and notches at different frequencies in the echo spectrum (Fig. 8D) into an estimate of the delay separation required to create these peaks and notches by interference. It is not sufficient to know just these *frequency* values, however; deconvolution requires knowledge of the values for the *periods* of the frequencies corresponding to the tops of the peaks and the bottoms of the notches. To be complete, deconvolution also requires knowledge about the detailed shape of the ridges in the echo spectrogram in the vicinity of the peaks and notches. (This is the previously-mewntioned phase information inherent in the shape of the sweeps.) The frequencies of the spectral peaks (f_p) are related to the reciprocal of the time separation of the overlapping echoes (ḍt):

$$f_p = n/ḍt \quad (\text{where } n = 1,2,3...)$$

Similarly, the frequencies of the notches (f_n) are related to the echo time separation (_t):

$$f_n = (2n+1)/2ḍt \quad (\text{where } n = 0,1,2,...)$$

For example, in Figure 8A-B, where ḍt = 60 µs, the spectral peaks (Fig. 8D) fall at 17, 33, 50, 66, 83, and 100 kHz (even frequency intervals), while the notches fall at 8.3, 25, 42, 58, 75, and 92 kHz (odd frequency intervals. The frequency spacing of the peaks and the notches is the reciprocal of the time separation ($df = 1/ḍt$) itself. Moreover, the even-frequency or odd-frequency spacing of these peak or notch frequencies specifies whether there is a phase-shift accompanying the time separation, as when one glint returns an echo that is 0° or 180° relative to the other echo. (Objects in air are so discontinuous from the air itself in acoustic impedance that the echoes they return are usually either 180° or 0° relative to the incident sound.)

4.4.2 Basis Vectors

The most complete implementation of the frequency-to-period knowledge required for deconvolution is reconstruction of the waveform of echoes at each frequency within the integration-time window for convolution. This can be achieved even after spectrograms have been formed (that is, after convolution) by using the spectrogram delays at individual frequencies (t_{A1}-t_{A5} in Fig. 8C) as time-

marking events to trigger the start of oscillatory signals, or *basis vectors*, that represent the original echo frequencies themselves. That is, each delay-line is used to register the arrival-time of one specific frequency in the echo, and the moment a coincidence between the echo and the emission at that frequency is detected, the basis vector begins to oscillate. In the SCAT model, the basis vectors (Fig. 8E) are cosine functions with durations sufficient to cover the interval of time acoss which the glint structure of echoes is to be reconstructed (usually the integration-time; but see scaling factor, below). However, the model is robust and works reasonably well for any periodic function used as basis vectors, even square-waves (Saillant, et al. 1993). The horizontal slices in the spectrograms (Fig. 8B) correspond to the frequency channels of the SCAT model, and each channel is tuned to a specific frequency in the emission or echo. Each channel then produces its own basis vector at a frequency that matches (or is proportional to) its original tuned frequency (Fig. 8E). The amplitude of the basis vector in each frequency channel is scaled according to the shape of the spectrum for the echo (from Fig. 8D to Fig. 8E). Two effects are achieved by using the spectrogram delays in different channels to start these oscillations in "cosine phase" separately for each channel. First, the slopes of the FM sweeps in the harmonics (FM_1,FM_2) are removed from subsequent concern by "dechirping" the signals. This permits the transmitted signals to be changed in duration and sweep-shape without having to keep track of this change. (One SCAT receiver processes all FM signals.) Second, information about changes in the shape of the sweep, or the phase of the different frequencies in echoes relative to emissions, is retained in differences in the starting time across the basis vectors. These starting times vary by up to one full period at each basis-vector frequency (compare basis vectors with each other in Fig. 8E).

4.5 Formation of SCAT Images

4.5.1 Reconstruction of the Glint Structure in Echoes

Once the basis vectors begin to oscillate, the next stage in the imaging process is simple: The arrival-times of the overlapping echoes (A and B) are reconstructed by summing the basis vectors across all 81 parallel channels (Fig. 8F) to form an average basis waveform. This average waveform is the image of the target's glint structure along the axis of echo-delay or target range. Due to reinforcement and cancellation of peaks and troughs in the basis vectors across channels, the original arrival-times of the echoes (t_A and t_B) appear as positive-going peaks in the resulting image even though no correspondingly well-resolved pair of events in the original echoes registers their arrival-time separation. The locations of the tips of the peaks can be estimated with considerable accuracy, depending upon the signal-to-noise ratio for the echoes. (Note that the SCAT model's reconstructed image in Fig. 8F resembles the shape of the images perceived by the bat in Fig. 7.) Because the frequencies, amplitudes, and phases of the basis vectors are controlled by the frequencies, amplitudes, and phases of the corresponding frequencies in the echo relative to the emission, the internal temporal organization of the echo can be reconstituted by itself. All other factors--FM sweep shape, harmonic structure, propagation delay to the target---are removed to reveal the contribution of the target in isolation. Moreover, because the basis vectors are aligned to start oscillating in cosine phase at the echo-delay values specified by the delay-lines, the entire image is displayed in absolute units of echo-delay or target range.

4.5.2 Time-Stretching of Basis Vectors and Images

One particularly significant feature of the SCAT model is the capacity to extract estimates of spectrogram delay (for echoes $A + B$ together) and delay separation (from echo A to echo B) using processing elements that have different time scales. It is convenient to introduce this feature by equating the frequency of each basis vector with the center frequency in each frequency channel of the spectrogram. In this case, the reconstructed image (Fig. 8F) has the same time-scale as the original signals; that is, the time between the first and second delay estimates is 60 μs in the time-series signal

formed by summation of the basis vectors, just as it was 60 μs in the original ultrasonic echoes (ḍt in the *echoes* [Fig. 8A-B] equals ḍt in the *image* [Fig. 8F]). However, this requires the basis vectors to be oscillations at ultrasonic frequencies, which is physiologically implausible. Oscillatory responses observed in the mammalian auditory system typically have frequencies of 100 Hz to about 1-2 kHz (Langner 1992; Langner and Schreiner 1988).

An alternative is to scale the frequencies of the basis vectors to be lower than the original ultrasonic signals, keeping the frequencies of the basis-vector oscillations *proportional* to the center frequencies of the band-pass filters rather than equal to them (Saillant et al. 1993). In this case, the ultrasonic frequencies in emissions and echoes extend hyperbolically from 20 to 100 kHz in Figure 8B, while the frequencies of the basis vectors extend hyperbolically from some *fraction* times 20 to 100 kHz in Figure 8E. This fraction is a *scaling factor* for the frequencies of the basis vectors; it lowers the frequencies in the reconstructed image and lengthens the spacing of the image components. That is, the time interval between the ultrasonic echoes (ḍt in Fig. 8A-B) is 60 μs while the corresponding time-interval between the first and second delay estimates in the time-series signal created by adding the basis vectors together (ḍt in Fig. 8F) could be 600 μs (for a scaling factor of 1/10) or 6 ms (for a scaling factor of 1/100). In the example in Figure 8, the original echo-delay separation is 60 μs, while the separation of the peaks in the reconstructed image is about 2.3 ms, so the scale factor is about 1/38. These longer time-intervals might realistically be represented by the timing of successive neural responses whereas the original 60-μs interval could not be represented directly by two successive neural responses only 60 μs apart.

5. Neural Responses in the Bat's Auditory System

The previous sections describe the acoustic stimuli received by echolocating bats, the images they perceive, and the computational requirements for creating these images from an initial representation consisting of hyperbolically-scaled spectrograms with an integration-time of about 350 μs. The critical feature of these images is their display of the arrival-times of both replicas of the sonar signal contained in overlapping echoes at time separations (ḍt) substantially smaller than this integration-time. We now turn to the problem of whether responses in the bat's auditory system manifest properties that are consistent with the requirements of deconvolution by the SCAT process or some near equivalent to it.

5.1 Principal Auditory Centers in the Bat's Brain

5.1.1 The Auditory Pathways

The auditory system of echolocating bats is much like the auditory systems of other mammals, except that is has been adapted to serve as a sonar receiver (Haplea, Covey, and Casseday 1994; Henson 1970; Pollak and Casseday 1989; Suga 1988, 1990). Figure 9 is a diagram showing the principal routes taken by neural responses to sounds as they follow the auditory pathways ascending from the cochlea through the bat's central nervous system (redrawn from Schweizer 1981). This diagram shows the principal auditory tracts leading from the *auditory nerve* of the bat's left ear to the major auditory nuclei depicted in cross-sections at four levels of the brain. The principal brain sites for processing acoustic information depicted in Figure 9 are the (1) the *cochlear nucleus* (designated *CN* in figures), which is the first synaptic step in auditory processing beyond the periphery; (2) a group of small nuclei (*trapezoid nuclei, medial and lateral superior olivary nuclei*) located along the ventral surface of the brain-stem, (3) the *nucleus of the lateral lemniscus* (*NLL*), which receives the output from the cochlear nucleus and other brain-stem sites through the lateral lemniscus, a large fiber tract connecting the brain-stem with the midbrain; (4) the *inferior colliculus* (*IC*), which is the major midbrain auditory center and a much-enlarged structure in bats, (5) the *medial geniculate*, which lies in

the thalamus just below the cerebral cortex, and (6) the *auditory cortex* (*AC*), which is usually described as the highest level of auditory processing and the presumed site of the physiological events that actually cause perception to happen.

5.1.2 Bilateral Connections of Auditory Centers

The diagram in Fig. 9 shows only the main excitatory projections from the *left* ear, which enter the left cochlear nucleus and then cross to the *right* side of the brain. This arrangement of crossed connections appears to make each auditory nucleus have its input from the *contralateral* ear. Information from the *ipsilateral* ear projects to most sites as well, however. Most structures above the level of the cochlear nucleus receive both contralateral and ipsilateral inputs, and the responses of their neurons exhibit varying degrees of *binaural* interactions. In the inferior colliculus of *Eptesicus*, for example, about 75% of neural responses recorded from single cells involve contralateral excitation and ipsilateral inhibition, often in combination with facilitation at some stimulus levels (Haresign et al. in prep).

5.2 Frequency Tuning of Neural Responses

The most pervasive feature of auditory coding in mammals is the tuning of neural responses at all levels of the auditory system to specific frequencies in sounds (for bats, see Henson 1970; Pollak and Casseday 1989; Suga 1988, 1990). Figure 10 illustrates tuning curves recorded from four levels in the auditory system of *Eptesicus* (see Fig. 9). All of the neurons at these sites respond selectively to frequencies used for echolocation (see Fig. 2).

5.2.1 Auditory Nerve, Cochlear Nucleus, and Nucleus of the Lateral Lemniscus

Figure 10A shows tuning for five representative cells from the anteroventral cochlear nucleus (AVCN), and Figure 10B shows five neurons from the posteroventral cochlear nucleus (PVCN). From what is presently known, the tuning curves from the AVCN can be taken as most representative of the frequency selectivity of primary auditory neurons at the input of the bat's auditory nervous system. Figure 10C-E shows tuning curves for representative cells from the two enlarged monaural divisions of the nucleus of the lateral lemniscus in *Eptesicus* (intermediate nucleus of the lateral lemniscus--INLL, ventral nucleus of the lateral lemniscus--VNLLm, VNLLc; Covey and Casseday 1991). Most of these tuning curves are *V-shaped*, with a sharp tip and progressively wider tuned regions above the tip. In contrast, many cells in the VNLLc (Fig. 10E) instead have very broad tuning curves. These broadly-tuned neurons nevertheless are capable of conveying frequency-specific information about echoes because their discharges register the entry of an FM sweep into the tuning curve and can also register amplitude modulations spread across the FM sweep as a whole (Covey and Casseday 1991).

5.2.2 Sharpness of Tuning at the Periphery

Figure 11 shows sharpness of tuning expressed as Q_{10dB} for cells in the cochlear nucleus of *Eptesicus* (Haplea et al. 1994). These Q_{10dB} values range from about 3 to 15 at frequencies from 10 to 70 kHz. The sharpness of tuning for first-order auditory neurons has been predicted from the rate-of-sweep at different frequencies in the bat's sonar sounds, on the assumption that the neurons will be tuned to optimize the accuracy of registering echo delay (Menne 198X). These predicted Q_{10dB} values, based on over 400 recorded echolocation sounds, are shown by dashed lines in Figure 11 (*mean ±1 standard deviation*). They fall in the same range as the measured Q_{10dB} values at those frequencies where tuning has been measured. Figure 11 also has a sloping line showing values of Q_{10dB} measured from frequency-response curves of the bandpass filters (*SCAT filters*; Saillant et al. 1993) used in the SCAT model of echolocation that provides a conceptual framework for this chapter. These filters have a frequency selectivity comparable to tuning curves in the bat, at least for the range of frequencies where Q_{10dB} values have been obtained.

5.2.3 Inferior Colliculus

Figure 10F-I shows tuning curves from the inferior colliculus of *Eptesicus* (Casseday and Covey 1992; Ferragamo, Haresign, and Simmons in press; Jen and Schlegel 1982; Poon et al. 1990). Tuning curves appear narrower in the inferior colliculus (Fig. 10F-I) because they have different shapes, with steeper, more nearly vertical, high-frequency and low-frequency slopes. A significant proportion of cells in the inferior colliculus also have "closed" tuning curves, with an upper limit to the response area (Fig. 10I). Echoes typically reach the bat's ears at amplitudes that are lower than the amplitudes of the transmissions picked up directly at the ears. Many cells with upper limits to their tuning curves would be unresponsive to the loud outgoing sound and thus would be selective for responding to *echoes* rather than transmissions. In the band of frequencies from 22 to 30 kHz, some of the tuning curves in the inferior colliculus of *Eptesicus* (Fig. 10H) appear narrower than those at 22 to 30 kHz in more peripheral nuclei. These especially sharply-tuned cells (called *filter* neurons; Casseday and Covey 1992) provide a narrow segment of the frequency axis from 22 to 30 kHz with an exaggerated sharpness of tuning probably produced by juxtaposition of excitatory and inhibitory inputs at slightly different frequencies (Casseday and Covey 1992; Suga and Schlegel 1973).

5.2.4 Auditory Cortex

Figure 10J illustrates tuning curves recorded from neurons in the auditory cortex of *Eptesicus* (Jen and Schlegel 1982). In anesthetized bats, cortical cells readily respond to tone-bursts, are tuned to a specific frequency, and have V-shaped tuning curves resembling those found in the cochlear nucleus (Fig. 10A-B) or nucleus of the lateral lemniscus (Fig. 10C-D). However, in awake bats, most cortical neurons are relatively unresponsive to tone-bursts; they require instead combinations of different frequencies and specific timing of multiple sounds to evoke an appreciable response. For example, neurons in a sizable population from the cortex of *Eptesicus* are tuned to more than one narrow frequency region (Dear et al. 1993; Fritz in press). Figure 10K illustrates a typical *multipeaked* tuning curve, with tuned frequencies centered at 15 kHz and 45 Khz, and fairly level-tolerant tuning curve around each tuned frequency. These cells mostly have a lower tuned frequency in the range of 10 to 40 kHz and a higher tuned frequency in the range of 30 to 80-90 kHz. Multipeaked cells are relatively unresponsive to frequencies in the interval between their tuned frequencies. A few cells with multiple tuned frequencies are found in the inferior colliculus of *Eptesicus* (Casseday and Covey 1992), but the multipeaked cells in the cortex are more common. Figure 12 shows the distribution of frequency ratios (f_2/f_1) for the high-frequency tuned region (f_2) to the low-frequency tuned region (f_2) from cortical multipeaked neurons in *Eptesicus*. A large proportion of the multipeaked cells have their two tuned frequencies in a frequency ratio around 2:1 or 3:1, with a smaller proportion of cells having intermediate ratios. The equations in Section 4.4.1 suggest that these multipeaked responses may embody frequency-domain information--"knowledge"--about the locations of spectral features necessary for deconvolution of overlapping echoes. The frequency spacings of peaks and notches fall at different frequency ratios that actually appear to be represented physiologically.

5.3 Distribution of Tuned Frequencies

5.3.1 Density of Frequency Tuning

Figure 13 illustrates the distribution of tuned frequencies for neurons at the cochlear nucleus (Haplea et al. 1994), nucleus of the lateral lemniscus (Covey and Casseday 1991), inferior colliculus (Casseday and Covey 1992; Ferragamo, Haresign, and Simmons in press; Jen and Schlegel 1982; Poon et al. 1990), and auditory cortex (Dear et al. 1993). These histograms show the emphasis placed on encoding information at different frequencies, with an especially large proportion of cells tuned to 20-40 kHz and a secondary proportion tuned to about 60-70 kHz.

Figure 14A replots the density of frequency tuning in the inferior colliculus with histogram bin-widths of 1 kHz from the 2-kHz bin-width shown in Figure 13C. This density distribution has two

segments, one with a peak at 25-30 kHz and steady decline to 55-60 kHz, and the other with a peak at 60-63 kHz and a decline to 95 kHz. (Neurons tuned to frequencies of 22 to 30 kHz include the sharply-tuned "filter" neurons in Fig. 10H.) The data shown along the horizontal frequency axis in Figure 14A are replotted along a period scale in Figure 14B by taking the reciprocal of the frequency value for each bin in Figure 14A. Figure 14B shows two jagged curves, with a break at 16.7 μs (the reciprocal of 60 kHz, which is approximately the location of the break between the two parts of the distribution in Fig. 14A). The curves in Figure 14B thus show a natural segmentation into two parts at about 16.7 μs. In Figure 14B, separate regression lines (*a*,*b*) are plotted for the short-period segment (12 μs to 16.7 μs) and the long-period segment (16.7 μs to 40 μs). When these two regression lines are transposed back onto the frequency scale of Figure 14A (curves *a* and *b*), they outline the shape of the original density distribution along the frequency axis. The shape of the regression curves in Figure 14A is hyperbolic due to their reciprocal relation to the straight regression lines in Figure 14B. Because the bat's sonar sounds have FM sweeps that are approximately hyperbolic in shape (Fig. 2), these curves trace the *dwell-time* of the bat's sonar sounds at each frequency for FM_1 (*a*) and FM_2 (*b*) (Fig. 2A-D).

5.3.2 Overrepresentation of Low Frequencies

In Figure 13 and 14A, frequencies of about 20 to 40 kHz are overrepresented (Casseday and Covey 1992) by neurons tuned to these frequencies compared to frequencies of 40 to 100 kHz. One aspect of this overrepresentation is the presence of filter neurons tuned to 25-30 kHz with especially sharp tuning curves (Fig. 10H). (*Eptesicus* probably uses these neurons to enhance detection for echoes of the relatively shallow FM sweeps in the range of 28-25 kHz that it broadcasts when searching for prey in open areas.) Approximately 24% of the neurons tuned to frequencies of 22 to 30 kHz are identified as filter neurons (Casseday and Covey 1992). However, even if the heights of the histogram bars in Figure 14A at 22-30 kHz are reduced by 24% to remove these specialized filter cells from the density distribution, the overrepresentation of frequencies in the 20-40 kHz region still exists because these histogram bars remain substantially higher than those at other frequencies.

The regression curves (*a*,*b*) in Figure 14A suggest that the density of frequency tuning might be numerically related to the period of each tuned frequency rather than directly to the frequency values themselves. To test this possibility, the entire data-set is transformed from the distribution of cells at each tuned frequency (1-kHz bins) in Figure 14A to the distribution of cells at each tuned period (1-μs bins) and plotted yet again in Figure 14C. Now the histogram density is more nearly uniform across different bins. In Figure 14C, the proportion of cells tuned to each period from 12-13 μs to about 35 μs varies by only a factor of roughly *two* from one histogram bar to another, whereas the proportion of cells tuned to each frequency in Figure 14A varies by a factor of as much as *eight* from one bar to another. Furthermore, the nonuniformity remaining in the distribution of Figure 14C is chiefly a higher proportion of cells tuned to periods of 33-38 μs compared to shorter periods. Periods of 33-38 μs correspond to frequencies of 26-30 kHz, which are frequencies to which the specialized filter neurons are tuned. If the heights of the histogram bars in Figure 14C at 33-38 μs are reduced by 24% to remove the proportion of filter neurons from the data, then the density of period tuning does appear to be approximately uniform across all periods from 12-13 μs to 35-40 μs. Thus, there is a genuine overrepresentation created by the presence of specialized filter neurons at low frequencies combined with a gradual skewing of the distribution towards lower frequencies as a consequence of the nearly uniform representation of ultrasonic periods across "nonfilter" neurons.

5.4 Tonotopic Organization of Tuned Frequencies

5.4.1 Nucleus of the Lateral Lemniscus

Figure 15A shows the mapping of frequencies along the frequency axis for one region of the nucleus of the lateral lemniscus in *Eptesicus* (VNLLc; Covey and Casseday 1986). This plot yields a

curvilinear relation from about 20 to 80 kHz. The same data are replotted in Figure 15B in terms of the period corresponding to each tuned frequency (Simmons et al. 1990). This relation is a straight line which can then be replotted back on frequency coordinates in Figure 15A as a regression curve. The tonotopic axis in the nucleus of the lateral lemniscus (VNLLc) of *Eptesicus* appears approximately hyperbolic with frequency, or linear with period.

5.4.2 Inferior Colliculus

Figure 15C shows a tonotopic axis for the inferior colliculus of *Eptesicus* taken from a rough three-dimensional reconstruction of its frequency-tuned layers (Poon et al. 1990). This relation between frequency and volume is curvilinear from about 20 to 80 kHz. Figure 15D shows the same volumetric relation in in terms of period, and once again, this topographic relation is approximately linear. Figure 15E-F shows similar relations for frequency and period from a more detailed set of measurements than in Figure 15C (Casseday and Covey 1992; Casseday, pers. comm.). The plot of period and volume in Figure 15F is as linear as the plot in Figure 15D. Thus, the tonotopic axis of the inferior colliculus in *Eptesicus* appears to be well-described as linear with period, or hyperbolic with frequency.

5.4.3 Auditory Cortex

Spatial representation of frequency is the most global characteristic of the auditory cortex in *Eptesicus* (Dear et al. 1993; Jen et al. 1989). However, the cortical frequency contours are quite convoluted, and individual bats differ in their tonotopic maps, which makes it difficult to come up with a composite map that reliably depicts all frequency regions. Frequency scales estimated from the two available tonotopic maps are shown in Figures 15G (Dear et al. 1993) and 15I (Jen et al. 1989). In both cases, there is considerable scatter in the data-points compared to the tonotopic axes for the nucleus of the lateral lemniscus (Fig. 15A) or the inferior colliculus (Fig. 15C,E). When these cortical data are replotted in terms of period (Fig. 15H,J), a straight regression line is about as accurate for characterizing the frequency plot as for characterizing the period plot. It not clear whether the cortical tonotopic map is incomplete (more auditory area awaits recording), or whether the auditory cortex simply does not have a tonotopic organization that matches precisely the organization found at lower centers. For example, the presence of multipeaked tuned neurons that are unique to the cortex (Fig. 10K and Fig. 12) may affect the frequency organization of the cortex in comparison with other sites containing neurons tuned to just one frequency region.

5.5 On-Responses to FM Stimuli

The bat's auditory system marks the time-of-occurrence of individual frequencies in the FM sweeps of sonar emissions and echoes with *on-responses* to these frequencies (Pollak et al. 1977; Pollak and Casseday 1989; Suga 1970). For example, in the inferior colliculus of *Eptesicus*, 93% of the recorded cells in one study respond to their tuned frequencies in a 2-ms artificial FM echolocation sound with just one discharge (Ferragamo, Haresign, and Simmons in press). Figure 16A-C illustrates latency (or PST) histograms of on-responses in three different single neurons to a 2-ms FM sweep that passes through each cell's tuned frequency. All three of these cells respond with an average of a single on-discharge to each occurrence of the FM stimulus; one discharge invariably occurs and there is no prominent second or third discharge to distort the histogram from its peaked shape. Moreover, cells in the inferior colliculus typically respond to *both the emission and the echo* unless their tuning curves have an upper threshold which can block responses to the emission (Fig. 10I) or unless they have recovery-times that exceed the delay of the echo. The first cell (Fig. 16A) responds to a frequency between 30 and 10 kHz with a latency of 12.5 ms and a standard deviation of 70 us. The second cell (Fig. 16B) responds to a frequency between 40 and 20 kHz with a latency of 16.8 ms and a standard deviation of 390 us. The third cell (Fig. 16C) responds to a frequency between 30 and 10 kHz with a latency of 21.6 ms and a standard deviation of about 6 ms. (The fourth latency histogram in Fig. 16D

shows the on-responses of several cells recorded together as a small multi-unit cluster. The properties of these local multi-unit responses will be examined below.) Similar observations of the magnitude and variability of latencies have been made in the FM bats *Myotis lucifugus* (Suga 1970) and *Tadarida brasiliensis* (Pollak et al. 1977).

Many neurons in the cochlear nucleus or the nucleus of the lateral lemniscus of *Eptesicus* will respond to a tone-burst with a *sustained* response that persists for the duration of the sound (Covey and Casseday 1991; Haplea, Covey, and Casseday 1994), and neurons in the inferior colliculus often respond selectively to the duration of a sound with a well-defined response at the end if the stimulus falls within the window of the cell's duration tuning (Casseday, Ehrlich, and Covey 1994). However, for short FM sounds comparable to echolocation signals broadcast when the bat is at distances of less than 1-2 m from a target (Fig. 2), the effective stimulus at each frequency is very short. The sound sweeps through its frequencies so rapidly that the dwell-time of the sweep in the vicinity of any particular frequency is typically a fraction of a millisecond. Consequently, the responses of most cells will be brief, too.

5.6 Latencies of On-Responses at Different Frequencies

Fig. 9 illustrates the major centers in the bat's auditory system and the order in which they are activated by the onset of a brief sound. Fig. 17 shows the time-of-occurrence, or *latency*, of the on-responses in neurons tuned to different ultrasonic frequencies at the cochlear nucleus (Haplea, Covey, and Casseday 1994), the nucleus of the lateral lemniscus (Covey and Casseday 1991), the inferior colliculus (Haplea, Covey, and Casseday 1994; Ferragamo, Haresign, and Simmons in press), and the auditory cortex (Dear et al. 1993) of *Eptesicus*.

5.6.1 Cochlear Nucleus and Nucleus of the Lateral Lemniscus

The cochlear nucleus is the first site to respond following activation of the auditory nerve, with discharges occurring at latencies of about 1 to 5 ms (*CN* in Fig. 17). All the cells in the cochlear nucleus of *Eptesicus* are tuned to specific ultrasonic frequencies (Fig. 10A,B), and most cells respond at a latency of about 2-3 ms to tone-bursts at their tuned frequencies, with a slightly greater spread of latencies at lower frequencies of 20-40 kHz than at higher frequencies of 50-90 kHz. In the nucleus of the lateral lemniscus, on-responses occur about 2 to 5 ms after the stimulus (*NLL* in Fig. 17). In addition, on-responses in some cells tuned to lower frequencies of 20-40 kHz have latencies as long as 8-12 ms. The majority of these responses fall 1-2 ms after responses in the cochlear nucleus, which reflects the synaptic delay and propagation-time that intervenes between these two centers (Fig. 9).

5.6.2 Inferior Colliculus

In the inferior colliculus, the first on-responses occur at latencies of 3-6 ms (*IC* in Fig. 17). These starting latencies are roughly 1-2 ms longer than at the nucleus of the lateral lemniscus, as would be expected from intervening synaptic and conduction delays (Fig. 9). Latencies of 3-6 ms are only the shortest latencies at each frequency, however. The responses of the inferior colliculus are characterized principally by a wide *dispersion* of latencies from their minimum values of 3-6 ms out to as much as 50 ms (Jen and Schlegel 1982; Poon et al. 1990; Kuwabara and Suga 1993; Pollak and Casseday 1989). In Figure 17 (*IC*), latencies for responses in cells tuned to 20-45 kHz are the most widely spread, being densely distributed from 3-6 ms to 25 ms, and more sparsely distributed to as much as 30-40 ms. At frequencies above 45 kHz, responses in the inferior colliculus have latencies mostly from 3 to about 12 ms, with a sparse distribution to as much as 25-30 ms. These dispersed responses are all on-responses to a brief sound, not sustained responses to a long-duration sound. The occurrence of these delayed responses appears to be regulated by prolonged intervals of inhibition initiated by the stimulus followed by abrupt excitation (Casseday, Ehrlich, and Covey 1994; Park and Pollak 1993; Suga 1964). The time

at which this discharge occurs may be considerably delayed in relation to the stimulus (Fig. 17 *IC*), but it nevertheless is an on-response to the sound.

Figure 17 (*IC*) shows a strong overrepresentation of dispersed latencies at low frequencies of 20 to 40 kHz, but these frequencies are themselves overrepresented chiefly because the scale of frequency is nearly uniform with period rather than frequency (Figure 14A-C). The frequency axis for the graphs in Figure 17 is linear, which emphasizes this overrepresentation. Figure 18 shows the distribution of on-response latencies for the inferior colliculus using a hyperbolically-scaled vertical frequency axis, which is more realistic (see above). When the low frequencies are spread out to show cells at equal increments of period, the distribution of latencies appears more broad and uniform. In particular, the sparse density of cells with responses at frequencies of 50-80 kHz in Figure 17 (*IC*) is compressed in Figure 18 to create a density comparable to that observed at 20 to 40 kHz.

5.6.3 Auditory Cortex

Neural responses in the auditory cortex mirror the pattern of latency-dispersal observed at the inferior colliculus. Cortical on-responses begin at a latency of about 6-8 ms and are distributed rather densely for latencies as long as 15-20 ms at most frequencies (*AC* in Fig. 17). In addition, response latencies at lower frequencies of 20-40 kHz extend more sparsely for 30-35 ms or more, while responses at higher frequencies are mostly finished by about 15-20 ms. The initial latencies in the cortex are about 3 ms longer than in the inferior colliculus, but the cortex is separated from the inferior colliculus by another auditory center (Fig. 9). The medial geniculate intervenes in the ascending pathway to add more synaptic and propagation delays, which accounts for the larger latency increment between the inferior colliculus and the auditory cortex in Figure 17 than between the earlier stages.

6. Neural Processing of Echo Delay

6.1 Target-Ranging Computations based on Latencies

Echolocating bats determine the distance to a target by measuring the time that separates the echo from the emission (Simmons 1973). The mechanism for displaying echo delay is a supreme example of a well-defined auditory computation being carried out using neuronal circuits that have been identified, at least in outline (Casseday and Covey 1992; Ferragamo, Haresign, and Simmons in press; Jen and Schlegel 1982; Kuwabara and Suga 1993; Pollak and Casseday 1989; Park and Pollak 1993; Suga 1988, 1990; Sullivan 1982). These circuits operate upon the latencies of on-responses at each frequency in the broadcast sound, followed by latencies of on-responses to each frequency in the echo.

6.2 Delay-Lines in the Inferior Colliculus

6.2.1 Latencies as Delay Taps in Physiological Delay-Lines

The principle that underlies target ranging by bats is to retard neural responses to the emission for a sufficiently long interval of time that the echo comes back and its responses start to occur simultaneously with those for the emission. The broad dispersal of latencies in the inferior colliculus creates these delays. Because the sonar of *Eptesicus* has an operating range of 5 m, responses to the emitted sound have to be delayed for up to about 33-35 ms to insure that some will still be occurring when the earliest responses to the echo take place at their shortest latencies of 3-5 ms. The neurons whose responses implement these delays each produce a single discharge to the emission or the echo at a characteristic latency (Fig 16A-C). That is, the delay-taps in the physiological implementation of the delay-lines are single cells tuned to the same frequency, but with different latencies. For example, the responses of neurons tuned to 30 kHz in Figure 18 act as a *delay-line for 30 kHz* in the FM sweep. Similar subpopulations of neurons act as delay-lines at other frequencies.

The physiological delay system in the bat is not as simple as a set of delay-lines creating a long series of responses to the emission, with short-term responses to the echo. Most of the cells shown in

Figure 18 will respond to both the emission and the echo if the sounds are a few milliseconds apart, so each sound triggers a flurry of responses stretched out over 25-30 ms. The extended pattern of responses to the emission depicted in Figure 18 thus is followed by a similar extended pattern of responses to the echo. These patterns overlap each other at a time separation equal to the acoustic delay of the echo, and the coincidence comparisons that correlate the shape of the FM sweeps for the emission and the echo actually take place between these patterns of dispersed latencies rather than simply between the spectrograms. Computations for echo-delay based on coincidence-detection of dispersed latencies appear at first glance to be approximately equivalent to the *spectrogram correlation* stage of the SCAT model (Fig. 8A-C). However, the patterns of dispersed latencies themselves will presently be seen to have properties as time-series signals that go beyond mere delays to make coincidence-detection equivalent in some respects to the *spectrogram transformation* stage as well.

6.2.2 Accuracy of Time Registration by Delay-Taps

The latencies of on-responses in different neurons vary somewhat from one stimulus presentation to another, and this variability is assumed to limit the accuracy with which the arrival-time of echoes can be determined (see Haplea, Covey, and Casseday 1994; Pollak et al. 1977; Schnitzler, Menne, and Hackbarth 1984). The widths of the peaks in the latency histograms shown in Figure 16A-C are typical of on-responses in the inferior colliculus of *Eptesicus*; some are as narrow as 50-100 μs (A), and many are as narrow as 300-500 μs (B), but many also are up to several milliseconds wide (C). Figure 19 shows the distribution of different latency variabilities (standard deviations for widths of histogram peaks) across absolute latencies (Ferragamo, Haresign, and Simmons in press). Responses that have absolute latencies anywhere from 3 ms up to 25-30 ms also have latency variabilities anywhere from about 50-100 μs to 4 ms, with a wider scattering of variabilities over 4 ms. Latency variabilities from 50-100 μs up to about 2 ms are the most densely-represented values across most absolute latencies. Surprisingly, the variability of response timing does not simply increase as latency increases in Figure 19, as might be expected if longer latencies leading up to responses in the inferior colliculus bring more opportunities for jitter in latency to accumulate. There are numbers of cells with latency variabilities as small as 50-100 μs at absolute latencies all across the range from 4 ms to 25 ms. This is a significant result because accurate registration of the timing of sounds depends upon retention of narrow latency variability across a wide span of absolute latencies in at least some of the cells.

6.3 Delay-Tuned Responses and Coincidence-Detection

6.3.1 Tuning Curves for Echo Delay

Neurons in the inferior colliculus act collectively like a system of multiple-tap delay-lines for storing the time-of-occurrence of each frequency in the broadcast sound or echo (*e.g., delay-line at 30 kHz* in Fig. 18). However, this delayed representation does not by itself create a *display* of echo delay, it just registers both sounds as events occurring at different times. The display is created by neurons that *compare* responses to emissions and echoes at the next level of processing--neurons that are specialized for responding only to echoes at certain delays. Figure 20A illustrates responses of three neurons in the auditory cortex of *Eptesicus* that are "tuned" to different values of echo delay (Dear, Simmons, and Fritz 1993). One of these cells responds most strongly to echoes at a *best delay* (BD) of 5 ms, the second cell has a best delay of 12 ms, while the third cell has a best delay of 20 ms. The first cell thus responds most strongly to a target located 86 cm away, the second cell responds most strongly to a target at 2.1 m, while the third cell responds to a target at 3.4 m. Delay-tuned cells in the auditory cortex of *Eptesicus* provide more-or-less continuous coverage of delays from 2 to 28 ms, which corresponds to target ranges of 34 cm to 4.8 m (Dear et al. 1993).

6.3.2 Detection of Coincidences between Responses to Emissions and Responses to Echoes

The neurons which compare echoes with emissions respond to the *coincidence* between a response generated at one latency in the inferior colliculus to the emission and a response generated at another, necessarily shorter, latency to the echo. Figure 20B shows a dot-raster plot for a series of responses in a delay-tuned neuron to repeated presentations of an emission and an echo at this cell's best delay of 10 ms. This particular neuron responds with a single discharge every time an echo arrives at a delay of 10 ms. Each response is characterized by a fixed latency of 16.2 ms after the emission and a latency of 6.2 ms after the echo. The difference between these two latencies is the 10-ms value of echo-delay associated with delay tuning. Coincidence-detecting cells are located in the medial geniculate (Fig. 9), where they receive two inputs from the inferior colliculus at different latencies and then send their response signaling a coincidence between these inputs onward to the auditory cortex. Cortical cells which receive the coincidence-registration as their input are tuned to a best delay corresponding to the latency difference for the inputs delivered to the coincidence-detectors by the inferior colliculus (Casseday and Covey 1992; Ferragamo, Haresign, and Simmons in press; Jen and Schlegel 1982; Poon et al. 1990; Kuwabara and Suga 1993; Pollak and Casseday 1989; Suga 1988, 1990; Sullivan 1982). Delay-tuned neurons also are found in the midbrain of *Eptesicus*, in a structure located between the inferior colliculus and the superior colliculus (Dear and Suga in press; Feng, Simmons, and Kick 1978).

6.3.3 Temporal Accuracy of Delay-Tuning versus *Accuracy of Coincidence Registration*

In the examples in Figure 20A, the sharpness of delay-tuning is proportional to best delay. The width of delay-tuning (at 50% of maximum response) is ±5 ms at a best delay of 10 ms, ±6-7 ms at a best delay of 12 ms, and ±10 ms at a best delay of 20 ms. In other cortical neurons, delay-tuning width does not increase but is constant across different best delays, which makes it proportionally sharper at long delays than short delays (Dear, Simmons, and Fritz 1993). Still, however, the degree of selectivity to delay for each cell is measured in *milliseconds*, while the bat's behavioral echo-delay acuity is a fraction of a *microsecond*, and in the limiting case 10-15 *nanoseconds* (see Section 3.5). At the input to delay-tuning, the variability in the latencies generated by the inferior colliculus is mostly in the region from hundreds of microseconds to several milliseconds (Fig. 19), which also is substantially greater than the bat's behavioral acuity. Delay-tuning curves thus have widths that correspond more to the latency variability of their inputs than to the bat's perceptual acuity for echo delay. If the bat uses its delay-tuned cells to perceive echo delay, why are there no cells that come within a factor of 5000 of the precision exhibited by the bat as a whole? This discrepancy between physiological and behavioral measures of timing accuracy is very large, which suggests that the tuning curves for echo delay are not the only higher-level representation of delay available to the bat.

Figure 20B reveals one aspect of delay coding that has greater precision than the width of the delay-tuning curve--the timing of the discharges in the delay-tuned cell. The neuron illustrated in Figure 20B has a best delay of 10 ms and a delay-tuning width of about ±3.7 ms, which is a typical of delay-tuned cells. However, the dot-raster plot of response latencies for repeated presentations of an echo at this cell's best delay of 10 ms is very stable from one presentation to the next. Latencies are tightly clustered within a standard deviation of 300 μs even though the width of this cell's tuning curve for delay is about ±3.7 ms. The latency of this cell's response thus betrays about 20 times more delay-coding accuracy than is evident in the tuning curve alone, so the transformation of temporal information about echo delay into place by delay-tuning evidently is not a complete transformation (Simmons and Dear 1991). Information about delay still is retained in the timing of the response after the selectivity of the response to delay has been set up.

6.3.4 Progression of Delay-Tuned Responses and Build-Up of Cortical Range Images

Perhaps the most interesting feature of cortical delay-tuning curves in *Eptesicus* is the *evolution* of delay-tuning across the interval of about 30-35 ms following the broadcast sound. (Recall that this

interval is the *epoch* for return of all useable echoes of the emitted sound from targets as far away as 5 m.) The inferior colliculus of *Eptesicus* supplies a wide range of latencies for its responses to both the emission and the echo (Fig. 18), and delay-tuned neurons compare different combinations of emission latencies and echo latencies. Figure 21 shows the time after the broadcast (or emission latency) at which cortical delay-tuned neurons respond for different values of best delay from 2 to 28 ms (or 34 cm to 4.8 m of target range). The horizontal time axis in Figure 21 traces the progression of delay-tuned responses throughout the epoch for echo reception. The first event in the epoch--reception of the broadcast by the bat's ears--releases a sequence of events which enables different delay-tuned neurons to respond if an echo happens to arrive near any particular cell's best delay. This sequence is regulated by the passage of on-responses to the emission along the delay-lines of the inferior colliculus (see Fig. 18).

As time progresses following the broadcast (horizontal axis in Figure 21), echoes are received (upward-sloping *solid line* labeled *echo delay*) and registered by appropriate delay-tuned neurons after a minimum latency of about 5-6 ms (upward-sloping *dashed line* offset to right by this minimum *echo latency*). A curious feature of the system of dispersed latencies for responses to the emission and the echo is that each echo *continues to be registered* by delay-tuned neurons tuned to its arrival-time long after it has been received, up to the length of the whole epoch at 28 ms. This is made possible by the long span of latencies for responses to the emission and to the echo--as long as responses to both sounds occur concurrently and feed into the coincidence-detecting system, there will be delay-tuned responses to register that echo. For example, at a point in time 15 ms after the emission, delay-tuned neurons with best delays from 2 to about 8-9 ms are set to respond if an echo arrives (vertical shaded bar labeled *range image @ 15 ms*). These neurons create an image depicting the ranges of targets out to about 1.5 m. At a time 30 ms after the broadcast, neurons with best delays from 2 to 25 ms are set to respond (vertical shaded bar labeled *range image @ 30 ms*), so that the range image now extends out to 4 m. A good way to visualize this evolving range image is to imagine the shaded bar starting at the left of Figure 21 and moving to the right as time passes. The bar becomes progressively taller because echoes from further away continue to come in and be registered. In the narrow region around 2-4 ms immediately following the broadcast the bar is short, but it grows to encompass the full span of 2 to 28 ms at the end of the epoch of time for reception of all echoes out to the maximum operating range. Beyond this longest practicable delay, responses to the emission cease--they "run off the end" of the delay-lines--and delay tuning disappears. The range image itself thus is *dynamic*, moving in the latency-delay space of Figure 21 as a kind of vertically-spreading "A-Scope" display that accumulates new targets as their echoes arrive. The persistence of registration of a target *in each cell* is only a fraction of a millisecond (the duration of a neural discharge), but the persistence of registration *across the population of cells* at each delay is nearly 30 ms. Then, the image vanishes, to be refreshed by broadcast of the next sonar signal.

7. Temporal Organization of Response Latencies in the Inferior Colliculus

7.1 Reconstruction of FM Spectrograms from Dispersed Latencies

The latency histograms in Fig. 16A-C illustrate on-responses to FM sounds in three different neurons from the inferior colliculus of *Eptesicus*. Over 90% of the individually-recorded neurons respond to a 2-ms FM sound with an average of just one discharge per stimulus at a particular latency (Ferragamo, Haresign, and Simmons in press). Figure 18 shows the spread of latencies at each frequency to be quite broad, especially at first-harmonic frequencies of 20-50 kHz, but there does not appear to be any special temporal pattern or organization to these latencies--just a noisy-looking continuous distribution made up of a large number of latencies distributed over the range from about 4-5 ms to 25-30 ms. Nevertheless, if the timing of each on-response reflects the timing of the specific

tuned frequency for each neuron (Fig. 10F-I) in the stimulus (Bodenhamer and Pollak 1981; see Pollak and Casseday 1989), then the serial order of responses to different frequencies in an FM sweep ought to be mirrored in the serial order of responses tuned to different frequencies. In other words, the spectrogram of the FM sweep should reside in the pattern of latencies across neurons tuned to different frequencies; it should, in fact, be present in the data in Figure 18, but it is concealed by the overriding effects of latency dispersion itself.

The difference in latency for responses to each neuron's tuned frequency presented as a tone-burst compared to the latency of the response to an FM sweep containing that tuned frequency, should equal the latency change related to the position of each frequency in the sweep. Figure 22 shows this latency subtraction for 41 neurons in the inferior colliculus of *Eptesicus*, with an added correction for the periodic pattern of latencies in the tone-burst responses relative to the more phasic responses to short-duration FM sounds. The stimulus for each neuron is a 2-ms downward FM sweep with a width of 20-25 kHz that contains that cell's tuned frequency. The cells shown in Figure 22 have tuned frequencies from 15 to 85 kHz and latency variabilities (standard deviations) smaller than 2 ms (Fig. 19) to keep them compatible with the sweep duration. The latencies for each neuron in Figure 22 fall close to the location of the tuned frequency in the FM sweep, with an underlying distribution that is about 800 μs wide (50% width). The pattern of response-times in Figure 22 essentially recovers the time-frequency sweep itself, indicating that the initial auditory spectrogram representation is still present in the system of dispersed responses in the inferior colliculus in spite of the latency dispersion being larger than the differences in latency at different frequencies along the 2-ms FM sweep. The 800-μs width of this reconstructed spectrogram (shown schematically as integration-time in Fig. 5) is about the same as the behaviorally-measured integration-time in *Eptesicus*, which is ±350 μs (Simmons et al. 1989).

7.2 Patterns of Dispersed Latencies as Time-Series Signals

7.2.1 Extracellular Single-Unit Recordings

Each frequency-and-latency data-point in Figure 18 corresponds to a single cell out of the large number that were studied--the graph is made by combining data from numerous different neurons individually recorded during daily recording sessions from recording sites distributed throughout the frequency layers of the inferior colliculus. For all practical purposes, the pooled single-unit data in Figure 18 are randomly sampled from the distribution of dispersed latencies actually present in the inferior colliculus. Only one neuron is recorded at a time; the electrode is moved until one cell's discharges are well-isolated electrically from the discharges of other cells in the immediate vicinity. After one cell has been recorded, the electrode is moved to pick up discharges of a new cell. The usual practice is to move the electrode through a minimum distance of 100 μm or so before starting to hunt for another single-unit to record. This precaution avoids accidental re-recording of the same cell, but it costs the chance of recording responses within the 100-μm zone immediately surrounding the first cell. If there is any temporal organization to the responses of cells grouped within this small zone, it will not be observed in single-unit data.

7.2.2 Multi-Unit Responses

When the recording electrode is "between cells," the signal it picks up consists of discharges from several neurons mixed together about equally rather than being dominated by discharges from a single, well-isolated cell. This type of record is a ***multi-unit response***. Figure 16D shows a multi-unit response analyzed by the same level-discriminator method as the single-unit responses in Figure 16A-C and displayed as a latency (or PST) histogram. In this case there are three neurons whose on-responses appear prominently at the same electrode site. The three main peaks in the histogram correspond to their individual on-response latencies. These peaks are separated by a fixed latency interval of 2-3 ms. Numerous multi-unit recordings analyzed with several methods confirm the impression given by the

histogram in Figure 16D that the local organization of response latencies is not random but structured into a periodic sequence of on-responses. At each recording site, the neurons encountered in multi-unit signals have latencies spaced roughly at equal intervals of 1-3 ms and extending over a span of latencies from as little as 4-6 ms to as much as 20-25 ms. The three latency values shown in Figure 16D could have been picked up in the course of single-cell recordings but not during the same recording session because the electrode would be moved too far after the first cell's responses are recorded. Then, the latencies would be treated the same as the latencies in Figure 16A-C; they would be plotted in Figure 18 as completely independent data-points with no suggestion that they occur together locally, as the multi-unit recordings reveal to be the case.

7.2.3 Local Averaged Multi-Unit Responses

The most efficient way to process recordings of multi-unit responses is to treat them as miniature evoked potentials and average them in synchrony with the stimulus. Figure 23A shows a multi-unit response processed with a spike-level discriminator and displayed as a latency histogram. The stimulus is a digitally-synthesized 2-ms FM sound that has the same sweep structure and envelope as the echolocation sounds of *Eptesicus*. Figure 23B shows the multi-unit response from the same electrode site processed by averaging as an analog signal. The conventional spike-detection response in Figure 23A contains several neural discharges at different latencies, most prominently at 6, 8, and 12 ms. There is also a good deal of noise because the level-triggering technique only follows events in the envelope of the recording and cannot cancel these out over repetitive sweeps. The peaks in the histogram in Figure 23A must be responses of different neurons because single-unit recordings demonstrate that neurons in the inferior colliculus usually respond with ony one discharge to a short-duration FM stimulus (see Fig. 16A-C). The averaged response in Figure 23B registers the same peaks at 6, 8, and 12 ms, and also it reveals more structure in the signal, notably at 14, 18, 20-22, and 24 ms. Although the histogram suggests the presence of these other peaks at longer latencies, the averaged response reveals them more sharply because the level of noise intrinsic to the averages is lower. Events that are not correlated with the sound tend to cancel out, whereas they accumulate in the histogram. Another advantage of the averaged signal is the presence of both positive-going and negative-going waves that delineate each other's latencies more effectively than do the exclusively positive-going peaks in the histogram.

7.2.4 Latency Dispersion in Multi-Unit Responses

Because the inferior colliculus is organized tonotopically, local multi-unit responses are tuned to a specific frequency corresponding to the site of the electrode's tip along the frequency axis of the inferior colliculus (Fig. 15C-E). Figure 24 shows a series of averaged multi-unit responses recorded from the inferior colliculus of *Eptesicus* at different recording depths in the same electrode track (in 100-μm steps from 200 to 1600 μm; scale at left in Fig. 24). Each recording depth has a tuned frequency, and the frequencies sampled in this particular electrode penetration extend from 24 kHz near the surface of the inferior colliculus (bottom trace in Fig. 24) to 76 kHz at the deepest site recorded (top trace). The stimulus for the responses in Figure 24 is a 2-ms FM sound with two harmonics to mimic a sonar *emission* of *Eptesicus*. The averaged responses vary according to the position of the electrode, or according to the tuned frequency of each site. They begin at a latency of 4-5 ms and continue as a series of peaks at intervals of about 0.5 to 3 ms with latencies of at least 12 to 15 ms. At lower frequencies of 24 to 40 kHz the responses are relatively long-lasting, with latencies up to 15-20 ms or more. At progressively higher frequencies the length of the responses--the span of latencies covered by the peaks--becomes progressively shorter. In general, the pattern of latencies at different frequencies mirrors the dispersion of latencies in single-unit recordings (see Fig. 17 *IC*), which is to be expected if the local multi-unit responses consist of clusters of neurons responding together, but at

different latencies relative to one another. Very few neurons in the inferior colliculus of *Eptesicus* discharge more than one time to the 2-ms FM stimulus, so the pattern of dispersed latencies in Figure 18 represents different latencies for each *cell*, and, correspondingly, each peak in the multi-unit response in Figure 24 represents the contribution of either one cell or a small group of cells with the same, fixed latency. For the most part, then, different peaks are caused by responses to the FM stimulus in different cells, not different latencies for sequential responses in the same cells.

7.2.5 Periodic Organization of Multi-Unit Responses

The principal information added by the multi-unit responses is the prevalence of a periodic organization to the dispersal of latencies at each site. The latency histograms in Figures 16D and 23A show multiple peaks in the responses separated by intervals of 0.5 to 3 ms, and the peaks in the averaged response in Figure 23B match the latencies of discharges in the corresponding histogram. Crucially, the latencies of the peaks in the averaged multi-unit responses change systematically across recording depths with different tuned frequencies. In the upper range of stimulus frequencies (52-76 kHz), the peaks slide to longer latencies as frequency decreases. For example, the peak with a latency of 6 ms in the response at 76 kHz (upper trace in Fig. 24) shifts gradually to about 7 ms at lower frequencies of 71-73 kHz, and the following peak with a latency of 7 ms at 76 kHz shifts progressively to nearly 10 ms as frequency decreases from 76 to 58 kHz. Taken together, multi-unit responses to higher ultrasonic frequencies in FM sounds consist of three to perhaps six or seven large, nearly parallel ridges roughly 2 ms apart that slope downward to the right in Figure 24, with some lower-amplitude narrower peaks spaced 0.5-1 ms apart. The large ridges are not *exactly* parallel, however--the size of the shift in latency across frequencies generally is larger for peaks at longer absolute latencies. Their slope changes because the spacing of the ridges opens up from 1-2 ms near the start to 2-3 ms near the end. In contrast, at lower stimulus frequencies (24-40 kHz), the responses still contain peaks separated by 1-3 ms, but the peaks do not shift consistently to longer latencies at progressively lower frequencies. These responses contain more peaks over a longer span of time, and most peaks clearly shift gradually from one trace to the next, but their relation to stimulus frequency is complicated--the latency lengthens as frequency decreases from 40 kHz to 29 kHz and then shortens as frequency decreases further to 24 kHz.

7.3 Echo-Delay Coding by Multi-Unit Responses

7.3.1 Multi-Unit Responses as Signal Representations

Taken as a whole, the multi-unit responses from the inferior colliculus constitute a two-dimensional system of time-series events having the dimensions of time (latency to each peak; horizontal axis in Fig. 24) and frequency (ultrasonic tuned frequency; vertical axis in Fig. 24) *with respect to the external ultrasonic stimulus*. These in fact are the time and frequency dimensions for the spectrogram of the sound (Fig. 5). The chief difference is that the spectrogram consists of a single peak (300-400 μs wide; the integration-time) at a specific time representing the moment at which one of the frequencies in the FM sweep occurs, whereas the system of multi-unit responses consists of a series of peaks at fixed latencies, each with a width of (intriguingly) several hundred microseconds. The responses in Figure 24 are made up synchronous on-responses to the FM sound (see Fig. 23), and these, in turn, contain a spectrogram representation embedded in their dispersed latencies, as shown in Figure 22 and implied by earlier studies of response latencies in the inferior colliculus of bats (Bodenhamer and Pollak 1981).

Each one of the local responses has a waveform (a series of peaks) with a specific frequency composition (equivalent to the spacing of their peaks) and phase structure (equivalent to latencies of peaks) that is internal to the response itself. What makes averaged multi-unit responses potentially so important for understanding echolocation is that their internal structure changes systematically with the external ultrasonic stimulus. As signals in their own right, they contain frequencies of about 300 Hz to

2 kHz, with a tendency to have lower frequencies at the end of the response than at the beginning because the spacing of the peaks grows larger at longer latencies (the slope of the ridges located to the right in Fig. 24 is more gradual than the slope of the ridges located to the left). Moreover, their phases and durations change systematically with the stimulus, too. If the stimulus occurs later, the responses shift uniformly to a later time, and if the stimulus consists of lower ultrasonic frequencies, the responses become longer. Because the internal horizontal and vertical dimensions of the responses carry information about the time and frequency dimensions of the external stimulus, they constitute a signal representation perhaps best described as an *expanded spectrogram* (Ferragamo, Haresign, and Simmons in press).

7.3.2 Similarity of Multi-Unit Responses to Emissions and Echoes

The averaged multi-unit responses in Figure 24 were evoked by repetitive presentations of an FM stimulus comparable to an echolocation sound used by *Eptesicus*. Figure 25 shows multi-unit responses collected at the same recording sites for a *pair* of FM sounds that simulate a 2-ms FM sonar broadcast followed 7 ms later by an echo of the same sound. This time-interval is equivalent to the echo delay for a target range of 1.2 m. The multi-unit responses to each of the two sounds in Figure 25 are similar to the responses evoked by the single FM sound in Figure 24, but the responses to the two sounds together overlap considerably. In Figure 25, both the "emission" and the "echo" evoke multi-unit responses consisting of sequences of peaks at fixed latencies, with an interval of 7 ms between the extended series of peaks representing the emission and the corresponding series of peaks representing the echo. Essentially the entire pattern of sloping ridges evoked by the first sound is repeated 7 ms later for the second sound. The responses to both the emission and the echo last for 5 to 25 ms at different recording sites, with longer sequences of peaks at locations tuned to lower ultrasonic frequencies of 24 to 40 kHz. Because the delay of the echo is only 7 ms, the responses overlap for an appreciable time at most sites, beginning at about 12 ms following the emission (5 ms following the echo) and continuing until about 18-20 ms following the emission (11-13 ms following the echo). At higher ultrasonic frequencies of 73-76 kHz the responses in Figure 25 do not overlap, however, because the response to each sound is shorter than the delay of the echo.

7.3.3 Overlap and "Interference" of Multi-Unit Responses

The averaged responses to the emission and the echo in Figure 25 necessarily will overlap to some degree for all echo delays shorter than 20-25 ms, which corresponds to target ranges up to about 4 m. By thinking of the multi-unit responses as signals with a specific waveform (frequency, phase, duration), we can describe the region of overlap between the emission and echo responses as a region of *interference between the response waveforms*. In Figure 25, at sites with tuned frequencies from 24 to 71 kHz, the peaks in the two sets of responses are intermingled for latencies from 12 to at least 18-20 ms following the emission, or from 5 to 11-13 ms following the echo. The peaks in the overlapping waveforms at each of these sites intersect sum together or cancel each other depending on the polarity of the emission response and the echo response at each point in time, creating a pattern of interference that changes according to the delay of the echo after the emission. This pattern of interference carries much detailed information about echo delay that is not evident from considering individual response-peaks one-at-a-time, only from viewing the whole responses as periodic signals.

7.3.4 Delay-Tuning as "Read-Out" of Interference Patterns in Multi-Unit Responses

If the delay of the echo is less than 25-30 ms, then there will still be single-unit responses to the emission taking place in the inferior colliculus (Fig. 18) when the earliest responses to the echo also occur--that is, when the on-responses to the emission and the echo overlap. We thus see that the "region of overlap" between the emission and echo responses has already been described from single-unit data in terms of delay-lines at different frequencies (Fig. 18) followed by a system of coincidence-

detecting neurons that create delay-tuned responses (Figure 20). The coincidence-detecting neurons respond to simultaneously-occurring long-latency responses to the emission and shorter-latency responses to the echo (Fig. 20)--in other words, to overlap of one particular component of the response to the emission with a correspondingly particular response to the echo (Fig. 21). Viewed in this manner, the delay-tuned neurons can be thought of as monitoring the conditions of overlap between responses to emissions and echoes, registering the simultaneity of small segments of these responses--segments that correspond to individual peaks in the local multi-unit responses to emissions and echoes (Fig. 25). Although delay-tuned neurons respond selectively to echoes at delays that equal the difference between their emission and echo latencies (Figs. 20-21), the crucial fact that the local responses are *periodic* is not yet part of the delay-line model. By incorporating the locally-manifested periodicity of these responses into the already-described delay-tuning scheme, the higher-level delay-tuned neurons become a mechanism for reading out the interference pattern between responses to the emission and to the echo. Each peak in the multi-unit responses corresponds to the latency of on-responses in a small number of neurons (Fig. 23). Moreover, the relative latencies of on-responses to emissions and echoes are systematically extracted from the dispersed latencies of the inferior colliculus by delay-tuned neurons in the auditory cortex (Fig. 21). The occurrence or nonoccurrence of delay-tuned responses at specific combinations of ultrasonic tuned frequencies in the emission and the echo (vertical axis of Fig. 25) and differences between emission and echo latencies (horizontal axis of Fig. 25) directly maps the shape of the overlapping multi-unit responses into the activity of the auditory cortex, where it is presumed that patterns of activity lead to the formation of perceived images.

8. Sensitivity of Multi-Unit Responses to Fine Echo Delay and Phase

8.1 Are Multi-Unit Responses Functionally Equivalent to Basis Vectors in the SCAT Model?

The averaged multi-unit responses are periodic signals originating in the inferior colliculus and containing a series of waves at approximately constant frequencies (allowing for their noisiness and for the progressive widening of the interval between successive peaks at longer latencies, which would appear as a decrease in frequency over the duration of the response--see Fig. 25). As such, they are suggestive of the basis vectors in the SCAT model. In the model, the basis vectors are cosine-phase oscillations, but in the bat they have to be "synthesized" from sequences of local, synchronized on-responses at discrete latencies. In all probability no individual neuron in the inferior colliculus actually oscillates at the frequency of the multi-unit response (about 300 Hz to 2 kHz); instead, the on-responses of different neurons are interleaved locally to appear as a succession of events at that frequency. The question is whether these responses serve a function comparable to the computations developed in the SCAT process.

In one respect the discharges making up multi-unit responses more-or-less exactly fulfill one of the functions identified in the SCAT model--that of the delay-lines for determining the spectrogram delays ($t_{A1\text{-}5}$ in Fig. 8C) from coincidences detected at specific delay-taps. This component of the SCAT model correlates the emission and echo spectrograms. The model places additional computational significance on the basis vectors that goes beyond just delays, however, by locking their starting phases to the occurrence of delay-tap coincidences in the various frequency channels. In effect, the phase of the basis vectors is a surrogate for the phase of the ultrasonic frequencies in the FM sweep of the echo relative to the emission. The larger aspect of this "phase" is simply similarity between the slope of the ridges in the spectrogram for the echo relative to the slope of the ridges in the spectrogram for the emission, and this is subsumed into the delay-lines and coincidence-detectors common to the SCAT model and the bat. However, a smaller but still critical part of this "phase" in the SCAT model is the detailed variation in the timing of coincidences in adjacent frequency channels. Small variations in phase across frequency channels are transposed into the basis vectors as variations in the starting-time

of the cosine oscillations (Fig. 8E), which in turn are critical for accurate reconstruction of the arrival-times of overlapping echoes (Fig. 8F). While the interference pattern produced by the overlapping multi-unit responses in Figure 25 seems related to the delay-lines and coincidence detectors in the SCAT model, the multi-unit responses also have to be sensitive to the phase and fine temporal structure of echoes if they are to qualify as candidates for biological basis vectors like those in the model. The astonishing thing is that these responses in fact *are* sensitive to the phase of echoes, even though the stimuli are at ultrasonic frequencies.

8.2 Phase-Coherent Multi-Unit Physiological Responses

8.2.1 Responses to Binaural Echo-Delay Differences

Figure 26 shows an example from an experiment illustrating the mode for representing echo phase and fine echo delay by multi-unit responses in the inferior colliculus of *Eptesicus*. First, the stimuli for these experiments are illustrated in Figure 26A-B. To obtain independent control of binaural stimuli, the sounds are produced with earphones inserted into the ear canal of the bat's left and right ears. Figure 26A shows the waveform of an electronically-delivered 2-ms FM sonar emission (pulse, *P*) and echo (*E*) at a delay of 6 ms (corresponding to a target range of about 1 m) delivered to the bat's ipsilateral ear (opposite to inferior colliculus from which the recording was made). Below this ipsilateral trace is a series of echoes delivered at the contralateral ear at slightly different delays around 6 ms. In successive contralateral traces the echo varies in delay from 25 μs *before* 6 ms to 25 μs *after* 6 ms. Each successive trace shows a change of 5 μs in this contralateral echo delay. The interaural time difference (ITD) for these binaurally-delivered echoes varies from -25 to +25 μs around the absolute delay of 6 ms. Binaural time differences of this magnitude are equivalent to differences in target azimuth from about 18° ipsilateral to about 18° contralateral. Next, Figure 26B shows expanded views of the contralateral echo waveforms in the region where the first-harmonic FM sweep passes through 25 kHz (note time scales of 1 ms in Fig. 26A and 100 μs in Fig. 26B). These "zoom" views of the stimulus show individual cycles of the echoes and graphically confirm that their peaks shift to the right in 5-μs steps as the binaural delay difference (ITD) changes. The electronically-produced echoes delivered to the ipsilateral and contralateral ears indeed do differ in their arrival-time by amounts from -25 μs to 25 μs.

Figure 26C shows local averaged multi-unit responses from the inferior colliculus of *Eptesicus* recorded for the stimuli in shown in Figure 26A. (This particular recording site is tuned to 20-22 kHz.) Each trace in Figure 26C is similar to one of the overlapping multi-unit responses shown in Figure 25 at sites tuned to 24-28 kHz, but now there is a new dimension to consider. (In Figure 26C the responses are stacked vertically to demonstrate how binaural echo-delay differences (ITDs) are manifested in the responses. Again, note the time scales of 10 ms in Fig. 26C and 1 ms in Fig. 26D.) The peaks in these responses shift to longer latencies as the binaural echo-delay difference changes from -30 μs to 30 μs in steps of 5 μs. This range of binaural delay differences corresponds to target azimuths of about 22° ipsilateral to 22° contralateral. The significant feature of this effect is that the time-shifts in the responses are many times larger than the binaural echo-delay changes themselves. Figure 26D shows an expanded section of the neural responses from Figure 26C to illustrate the magnitude of the response time-shift in relation to the original stimuli in Figure 26B. In this example, the amount of time expansion in the responses is by a factor of about 50; that is, a 5-μs change in binaural delay difference leads to a 250-μs change in response latency. Moreover, it is completely surprising that these responses appear in some respects to *reconstruct* the stimulus waveform on this stretched time scale (compare Fig. 26B with Fig. 26D). Notice also that the shift in response latency affects the entire series of peaks evoked by both the emitted pulse (*P*) and the echo (*E*) as a function of the binaural echo-delay difference (ITD). As the delay difference changes from -25 μs to +25 μs, all the peaks in the responses

slide to the right--that is, to longer latencies, even those peaks that we would initially regard as part of the response to the emission because they precede the arrival of the echo. Evidently the responses to the emission and the echo belong to a single system of events that are collectively altered just by changes in the delay of the echo, presumably because responses to one pulse-echo pair are modified by the previous echo.

8.2.2 Responses to Echo Phase Shifts

Not only do the latencies of local multi-unit responses in the inferior colliculus represent binaural echo-delay differences on an expanded time scale, but they also represent the *phase* of the echo waveform itself on an expanded time scale. Figure 26D shows segments of the neural responses to illustrate the effect on latencies of what amount to *binaural* phase changes. Figure 27A now illustrates the same response segment at a binaural echo-delay difference (ITD) of zero (target straight ahead) for two added conditions--0° and 180° echo phase-shifts relative to the phase of the simulated emission. Note how the local multi-unit response is both *shifted to the right* and *expanded* slightly just as a result of the 180° echo phase shift. Figure 27B goes even further by showing a series of neural responses for different binaural delay differences combined with echo phase shifts of 0° or 180°. The latencies of the response peaks represent *both* interaural echo-delay differences and the phase of the echo in both ears together! Furthermore, the size of the latency shift is much larger than the size of the temporal shift in features of the ultrasonic echoes, where the phase-shift amounts to a displacement of the cycles in the echo waveform of only 5 to 25 μs and the binaural difference is only -30 to +30 μs.

8.2.3 Multiple Time Scales in Multi-Unit Responses

The changes in response latency illustrated in Figures 26-27 reveal a new scheme of representation in the inferior colliculus. First, the time-intervals between responses to the emission and responses to the echo represent the delay of echoes on a time scale that closely matches real-time in the stimuli (Figure 25). That is, peaks in the multi-unit response to the echo lag the corresponding peaks in the multi-unit response to the emission by an interval approximately equal to echo delay. Thus, the echo-delay axis of an ordinary A-Scope display prevails along the horizontal time axis of the multi-unit responses. However, the entire *pattern of latencies* in the local responses also appears to represent details of the fine phase and delay structure of the echoes on a wholly different time scale. Shifts of a few microseconds in the binaural delay or the phase structure of the waveform of echoes lead to shifts of hundreds of microseconds in the multi-unit responses. This discovery leads to a very novel conclusion: The bat may be able to "read" information from these responses at time-scales quite different from that actually observed electrically. Time-scale magnifications of this sort may account for aspects of the bat's performance that until now have appeared impossible to reconcile with physiological results, including echo-delay acuity of 10-15 ns, two-point echo-delay resolution of about 2 μs, and sensitivity to changes in echo phase. While the latency axis appears to be associated with variabilities of hundreds of microseconds (the width of the response-peaks), the true variability in the latencies of the responses is only a few microseconds as far as the fine delay and phase of echoes is concerned. The bat's practice of packing multiple time axes into the same neural signals is, to say the least, unexpected from previous physiological considerations, although it is anticipated by the time-scaling feature of the basis vectors in the SCAT model.

9. Cortical Responses to Echo Waveforms

9.1 Binaural Echo-Delay Differences

It is widely presumed, but not, of course, demonstrated, that the content of perceived images is generated by neural activity taking place in the cerebral cortex. It therefore in important to know whether the unusual latency shifts and their associated time-expansions seen in the inferior colliculus (Fig. 26-27) also occur in cortical responses. Figure 28 shows averaged multi-unit responses recorded

from the auditory cortex of *Eptesicus* that confirm the presence of these time-domain events at the highest level of auditory representation. (The format for this figure is similar to that in Figs. 26-27.) The acoustic stimulus consists of a 2-ms FM emission and then an echo arriving 15 ms later. This recording shows a large peak (*) corresponding to the discharge of a single, well-isolated delay-tuned cell (tuned to a delay of 15 ms) accompanied by smaller peaks which most likely originate from other cells in the vicinity. The complex of peaks in this part of the response shifts in latency by relatively large amounts in response to small changes in the binaural delay difference (-50 μs to +50 μs) delivered around the overall echo delay of 15 ms. As in Figures 26 and 27, the magnitude of change in latency appears magnified with respect to the magnitude of the change in the original echo waveform. The expanded view of this latency shift at the bottom of Figure 28 shows how the magnification effect "rides on top" of what would otherwise just be considered an ordinary delay-tuned response from a cortical single unit (see Fig. 20). Not shown here is the effect of changing the phase of echoes relative to emissions by 0° or 180°; the latency of cortical multi-unit responses also changes by a large amount for echo phase-shifts, as already shown for the inferior colliculus in Figure 27. Both of these effects are consistent with a possible role for delay-tuned coincidence-detecting neurons in reading out information about the overlap of local multi-unit responses to the emission and echo (Fig. 25), but it is only an indication of what might be happening as a mechanism for echo-processing, not a convincing proof of basis vectors in the brain. (At this stage of our knowledge about the mechanisms of perception, even unconvincing evidence is a step forward.)

9.2 Delay-Separation of Overlapping Echoes

The chief concern of this chapter is the bat's ability to perceive the arrival-times of closely-spaced echoes (Fig. 7), and the last question to consider is whether the same magnification of time-scales seen in the latencies of multi-unit responses is also seen in response to changes in the delay-separation of closely-spaced echoes. Figure 29 illustrates this type of response: The graphs show a series of multi-unit responses evoked by FM stimuli that mimic a sonar emission (12-ms duration) and a two-glint echo at a delay of 24 ms. This series shows the effect of changing the delay of the second of the two overlapping components of the echo from 0 to 200 μs in steps of 25 μs while the first component remains fixed at a delay of 24 ms. The latency of the principal peaks in the response shifts by about 3 ms for a 175-μs change in separation of the echo components, which is a magnification factor of about 17. This value is well within the range of magnification factors seen in the inferior colliculus for echo phase and binaural delay differences. For all practical purposes, the responses in Figure 29 may be part of the bat's "A-Scope" dispaly of the second of the two overlapping echoes (at delay t_B in Fig. 8). Thus, as predicted at least in a loose way by the SCAT model, the big brown bat seems to encode the delay-separation of closely-spaced, overlapping echoes by the timing of neural responses. It presently is unclear how many cells contribute to the multi-unit responses or what mechanism creates the time-scale magnification, but the physiological reality of these effects is well-demonstrated by Figures 26-29.

10. Summary and Conclusions

10.1 The SCAT Model and the Bat

The SCAT model (*1*) represents the waveform of FM sonar emissions and echoes (Fig. 8A) as spectrograms with 81 parallel frequency channels (Fig. 8B), (*2*) determines the time-separation beween the emission and the echo by measuring spectrogram delays in each channel using delay-lines that register coincidences between delayed representations of the emission and immediate representations of the echo ($t_{A1\text{-}5}$ in Fig. 8C), (*3*) triggers the occurrence of oscillatory basis vectors in each channel from the coincidences detected in the delay-lines (Fig. 8D-E), and (*4*) creates an image of echo delay by

summing the basis vectors across all the channels (Fig. 8F). The multi-unit responses suggest that a similar process may underlie echolocation: Perhaps the bat (*1*) represents the waveform of FM sonar emissions and echoes as spectrograms with numerous parallel frequency-tuned receptors and sharpens the temporal registration of each frequency in the spectrogram with on-responses in lower auditory centers (*CN, NLL* in Fig. 17), (*2*) disperses this spectrogram representation by triggering sequences of responses over the course of at least 20-25 ms after the emission or the echo in the inferior colliculus (Figs. 18 and 22), (*3*) organizes these dispersed latencies at periodic intervals (Figs. 23 and 25), and (*4*) determines the time separation of emissions and echoes from coincidences of responses at different latencies using the dispersed latencies as the equivalent of delay-lines to create delay-tuned neurons (Fig. 21). The chief difference between the SCAT model and this hypothetical description of echolocation is that the model triggers the periodic basis vectors off the coincidences detected at specific delay taps in the delay-lines (which themselves were activated by the initial registration of each frequency in the spectrogram of the FM sweep), while the bat triggers the periodic responses of local groups of neurons off the initial registration of each frequency in the spectrogram of the FM sweep and then uses these periodic responses as delay-lines for subsequently detecting coincidences between emissions and echoes using delay-tuned neurons. Essentially, the order in which the delay-lines and the periodic oscillations are introduced into the computations may be reversed for the bat relative to the model.

10.2 Physiological Representation of Echo Waveforms

Physiological events in the inferior colliculus of *Eptesicus* do not consist just of lots of on-responses; they consist of bursts of neurally-simulated "oscillations" when viewed as local averaged potentials recorded at different electrode sites in the inferior colliculus (Figs. 24-25). The phase of these multiple-peaked responses encodes the arrival-time and phase of echoes with surprising precision in a format not previously considered (Figs. 26-27). The principal new finding in the physiological data is that the time scale of the bat's phase-coherent physiological responses to FM echoes is not only *real-time* but also an *expanded or zoom time scale* that stretches significant elements of the echo waveform as time-series signals in the brain. These neural responses are sensitive to both binaural echo-delay differences and also echo phase, and they appear easily able to account for the bat's otherwise puzzling ability to perceive echo phase and submicrosecond echo delay. These properties are repeated in multi-unit responses recorded from the auditory cortex, with the added observation that the delay separation of closely-spaced echoes is included in this representation, too (Figs. 28-29). The presence of both normal time scales and expanded time scales in responses evoked within the auditory cortex greatly strengthens the liklihood that these multiple time scales are involved in creating the perceived images. It thus appears as though the big brown bat may perceive "A-Scope" images of targets at different ranges on a coarse time scale, plus "Zoom A-Scope" images of target azimuth and shape on expanded, finer time scales. These possibilities have not emerged from recent physiological studies by themselves but instead with the guidance of the SCAT model to point out parameters of neural responses that deserve more attention as candidate coding dimensions. The utility of the SCAT model in offering the basis vectors as a conceptual guidepost has been especially valuable for breaking new experimental ground in understanding the auditory computations for echolocation.

All three principal ways of influencing the waveform of sonar echoes at the bat's ears--changing their phase, changing their arrival-times at the left and right ears, changing the time-separation of overlapping components (A and B in Fig. 6)--are manifested as large shifts in response latency. Some means for neurally representing these dimensions of neural responses seems necessary to explain the bat's performance in critical tasks (see Simmons et al. chapter in Popper and Fay in press), even though it is frequently thought that bats cannot perceive these features due to physiological limitations on the

speed and accuracy of neural discharges (Pollak 1988, 1993; Schnitzler Menne and Hackbarth 1985). It appears as though some property the colliculo-cortical system in *Eptesicus* can rerepresent--actually *reconstruct*--time-series acoustic information delivered to the cochlea as time-series information made up of neural responses. It is astonishing that this reconstruction incorporates a magnification of the time-scale of the original acoustic waveform. From the point-of-view of this chapter, the most significant feature of these results is that the neural responses register acoustic time-series information at lower frequencies of the order of 300 Hz to 2 kHz, which are reasonable rates for neurons to operate at (Langner 1992), rather than at ultrasonic frequencies of 20 to 100 kHz, which presumably are beyond the capacity of neurons to encode directly.

References

Altes RA (1980) Detection, estimation, and classification with spectrograms. J Acoust Soc Am 67:1232-1246

Altes RA (1984) Texture analysis with spectrograms. IEEE Trans Sonics-Ultrasonics SU-31:407-417

Beuter KJ (1980) A new concept of echo evaluation in the auditory system of bats. In: Busnel R-G, Fish JF (eds) Animal sonar systems. Plenum Press, New York, pp 747-761

Bodenhamer RD, Pollak GD (1981) Time and frequency domain processing in the inferior colliculus of echolocating bats. Hearing Res 5:317-355

Casseday JH, Covey E (1992) Frequency tuning properties of neurons in the inferior colliculus of an FM bat. J Comp Neurol 319:34-50

Casseday JH, Ehrlich D, Covey E. (1994) Neural tuning for sound duration: Role of inhibitory mechanisms in the inferior colliculus. Science 264:847-850

Covey E, Casseday JH (1986) Connectional basis for frequency representation in the nuclei of the lateral lemniscus of the bat, *Eptesicus fuscus*. J Neurosci 6:2926-2940

Covey E, Casseday JH (1991) The monaural nuclei of the lateral lemniscus in an echolocating bat: Parallel pathways for analyzing temporal features of sound. J Neurosci 11:3456-3470

Dear SP, Fritz J, Haresign T, Ferragamo M, Simmons JA (1993) Tonotopic and functional organization in the auditory cortex of the big brown bat, *Eptesicus fuscus*. J Neurophysiol 70:1988-2009

Dear SP, Simmons JA, Fritz J (1993) A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. Nature 364:620-623

Dear SP, Suga N (in press) Delay-tuned neurons in the midbrain of the big brown bat. J Neurophysiol

Feng AS, Simmons JA, Kick SA (1978) Echo detection and target-ranging neurons in the auditory system of the bat, *Eptesicus fuscus*. Science 202:645-648

Ferragamo M, Haresign T, Simmons JA (in press) Response properties in the inferior colliculus of the echolocating bat, *Eptesicus fuscus*: Frequency and latency dimensions of expanded neural spectrograms. J Comp Physiol A

Griffin DR (1958) Listening in the dark. Yale Univ. Press, New Haven CT (reprinted by Cornell University Press, Ithaca NY, 1986)

Grinnell AD (1963) The neurophysiology of audition in bats: Temporal parameters. J Physiol 167:67-96

Haplea S, Covey E, Casseday JH (1994) Frequency tuning and response latencies at three levels in the brainstem of the echolocating bat, *Eptesicus fuscus*. J Comp Physiol

Hartley DJ (1992) Stabilization of perceived echo amplitudes in echolocating bats: II. The acoustic behavior of the big brown bat, *Eptesicus fuscus*, while tracking moving prey. J Acoust Soc Am 91:1133-1149

Henson OW Jr (1970) The ear and audition. In: Wimsatt WA (ed) Biology of bats, vol 2. Academic Press, New York, pp 181-263

Jen PHS, Schlegel PA (1982) Auditory physiological properties of neurons in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. J Comp Physiol A 147:351-363

Jen PHS, Sun X, Lin PJJ (1989) Frequency and space representation in the primary auditory cortex of the frequency modulating bat *Eptesicus fuscus*. J Comp Physiol A 165:1-14

Kick SA (1982) Target detection by the echolocating bat, *Eptesicus fuscus*. J Comp Physiol 145:431-435

Kick SA, Simmons JA (1984) Automatic gain control in the bat's sonar receiver and the neuroethology of echolocation. J. Neurosci 4:2725-2737

Kober R, Schnitzler H-U (1990) Information in sonar echoes of fluttering insects available for echolocating bats. J Acoust Soc Am 87:874-881

Kurta A, Baker RH (1990) *Eptesicus fuscus*. Mamm Species 356:1-10

Kuwabara N, Suga N (1993) Delay lines and amplitude selectivity are created in subthalamic auditory nuclei: The brachium of the inferior colliculus of the mustached bat. J Neurophysiol 69:1713-1724

Langner G (1992) Periodicity coding in the auditory system. Hearing Res 60:115-142

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat: I. Neuronal mechanisms. J Neurophysiol 60:1799-1822

Lawrence BD, Simmons JA (1982) Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. J Acoust Soc Am 71:585-590

Licklider JCR (1951) A duplex theory of pitch perception. Experientia 7:128-134

Menne D (1985) Theoretical limits of time resolution in narrow band neurons. In: Michelsen A (ed) Time resolution in auditory systems. Springer-Verlag, New York, pp 96-107

Menne D (1988) Is the structure of bat echolocation calls an adaptation to the mammalian hearing system? J Acoust Soc Am 83:2447-2449

Menne D, Kaipf I, Wagner I, Ostwald J, Schnitzler HU (1989) Range estimation by echolocation in the bat *Eptesicus fuscus*: trading of phase versus time cues. J Acoust Soc Am 85:2642-2650

Moss CF, Schnitzler H-U (1989) Accuracy of target ranging in echolocating bats: Acoustic information processing. J Comp Physiol A 165:383-393

Moss CF, Simmons JA (1993) Acoustic image representation of a point target in the bat, *Eptesicus fuscus*: Evidence for sensitivity to echo phase in bat sonar. J Acoust Soc Am 93:1553-1562

Moss CF, Zagaeski M (1994) Acoustic information available to bats using frequency-modulated sounds for the perception of insect prey. J Acoust Soc Am 95:2745-2756

Neuweiler G (1990) Auditory adaptations for prey capture in echolocating bats. Physiol Rev 70:615-641

Novick A (1977) Acoustic orientation. In: Wimsatt WA (ed) Biology of bats, vol 3. Academic Press, New York, pp 73-287

Park TJ, Pollak GD (1993) GABA shapes a topographic organization of response latency in the mustache bat's inferior colliculus. J Neurosci 13:5172-5187

Pollak GD (1988) Time is traded for intensity in the bat's auditory system. Hearing Res 36:107-124

Pollak GD (1993) Some comments on the proposed perception of phase and nanosecond time disparities by echolocating bats. J Comp Physiol A 172:523-531

Pollak GD, Casseday JH (1989) The neural basis of echolocation in bats. Springer-Verlag, New York

Pollak GD, Marsh DS, Bodenhamer R, Souther A (1977) Characteristics of phasic on neurons in inferior colliculus of unanesthetized bats with observations relating to mechanisms for echo ranging. J Neurophysiol 40:926-942

Poon PWF, Sun X, Kamada T, Jen PHS (1990) Frequency and space representation in the inferior colliculus of the FM bat, *Eptesicus fuscus*. Exp Brain Res 79:83-91

Popper A, Fay RR (in press) Handbook of auditory research. Springer-Verlag, New York

Pye JD (1980) Echolocation signals and echoes in air. In: Busnel R-G, Fish JF (eds) Animal sonar systems. Plenum Press, New York, pp 309-353

Saillant PA, Simmons JA, Dear SP, McMullen TA (1993) A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: The spectrogram correlation and transformation receiver. J Acoust Soc Am 94:2691-2712

Schmidt S (1992) Perception of structured phantom targets in the echolocating bat, *Megaderma lyra*. J Acoust Soc Am 91:2203-2223

Schnitzler H-U, Henson OW Jr (1980) Performance of airborne animal sonar systems: I. Microchiroptera. In: Busnel R-G, Fish JF (eds) Animal sonar systems. Plenum Press, New York, pp 109-181

Schnitzler H-U, Menne D, Hackbarth H (1985) Range determination by measuring time delay in echolocating bats. In: Michelsen A (ed) Time resolution in auditory systems. Springer-Verlag, New York, pp 180-204

Schweizer H (1981) The connections of the inferior colliculus and the organization of the brainstem auditory system in the greater horseshoe bat (*Rhinolophus ferrumequinum*). J Comp Neurol 201:25-49

Simmons JA (1973) The resolution of target range by echolocating bats. J Acoust Soc Am 54:157-173

Simmons JA (1979) Perception of echo phase information in bat sonar. Science 207:1336-1338.

Simmons JA (1989) A view of the world through the bat's ear: The formation of acoustic images in echolocation. Cognition 33:155-199

Simmons JA (1992) Time-frequency transforms and images of targets in the sonar of bats. In: Bialek W (ed) Princeton lectures on biophysics. World Scientific Publishing, River Edge NJ, pp 291-319

Simmons JA (1993) Evidence for perception of fine echo delay and phase by the FM bat, *Eptesicus fuscus*. J Comp Physiol A 172:533-547

Simmons JA, Chen L (1989) The acoustic basis for target discrimination by FM echolocating bats. J Acoust Soc Am, 86:1333-1350

Simmons JA, Dear SP (1991) Computational representations of sonar images in bats. Curr Biol 1:174-176

Simmons JA, Grinnell AD (1988) The performance of echolocation: The acoustic images perceived by echolocating bats. In: Nachtigall P, Moore PWB (eds) Animal sonar: Processes and performance. New York, Plenum Press, pp 353-385

Simmons JA, Kick SA (1984) Physiological mechanisms for spatial filtering and image enhancement in the sonar of bats. Ann Rev Physiol 1984 46:599-614

Simmons JA, Freedman EG, Stevenson SB, Chen L, Wohlgenant TJ (1989) Clutter interference and the integration time of echoes in the echolocating bat, *Eptesicus fuscus*. J Acoust Soc Am 86:1318-1332

Simmons JA, Ferragamo M, Moss CF, Stevenson SB, Altes RA (1990) Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation. J Comp Physiol A 167:589-616

Simmons JA, Moss CF, Ferragamo M (1990) Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus*. J Comp Physiol A 166:449-470

Skolnik MI (1962) Introduction to radar systems. McGraw-Hill, New York.

Suga N (1964) Recovery cycles and responses to frequency modulated tone pulses in auditory neurons of echolocating bats. J Physiol 175:50-80

Suga N (1970) Echo-ranging neurons in the inferior colliculus of bats. Science 170:449-452

Suga N (1988) Auditory neuroethology and speech processing: Complex-sound processing by combination-sensitive neurons. In: Edelman GM, Gall WE, Cowan WM (eds) Auditory function. Wiley, New York, pp 679-720

Suga N (1990) Cortical computational maps for auditory imaging. Neural Networks 3:3-21

Suga N, Schlegel P (1973) Coding and processing in the nervous system of FM signal producing bats. J Acoust Soc Am 84:174-190

Sullivan WE (1982) Neural representation of target distance in auditory cortex of the echolocating bat *Myotis lucifugus*. J Neurophysiol 48:1011-1032

ILLUSTRATIONS

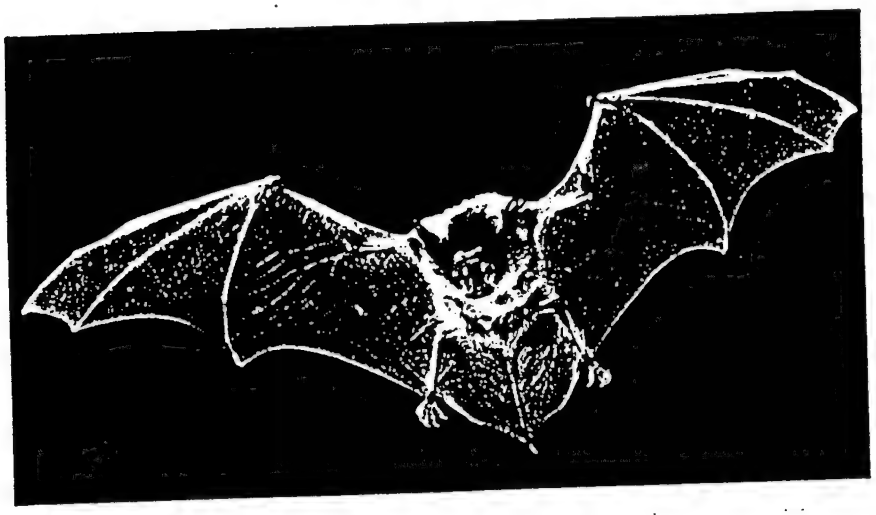

Fig. 1. The big brown bat, *Eptesicus fuscus*, approaching a target (photographed from the target's position by S. P. Dear and P. A. Saillant.)

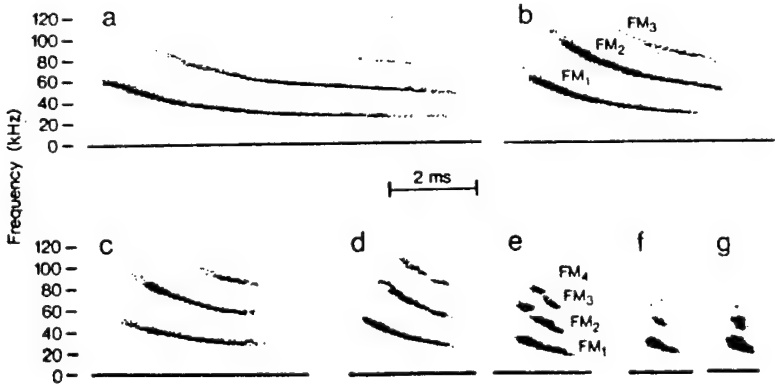

Fig. 2. Spectrograms of multiple-harmonic FM echolocation sounds emitted by *Eptesicus* at different stages of insect-pursuit and capture while being photographed and videotaped in laboratory studies of interception (Fig. 1). (a) search stage, (b,c,d) approach or tracking stage, (e,f,g) terminal stage (see Popper and Fay in press).

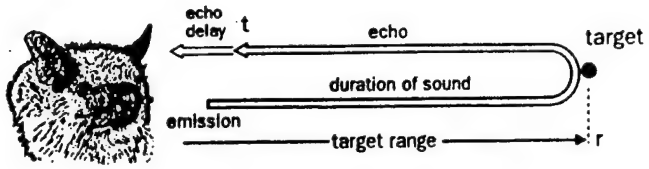

Fig. 3. Diagram showing duration of bat's sonar sound in relation to target range and echo delay.

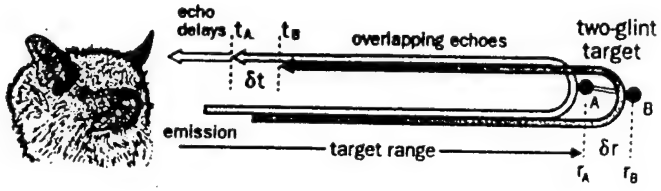

Fig. 4. Diagram showing duration of bat's sonar sound and overlapping echoes from two parts of the same dipole target at slightly different distances.

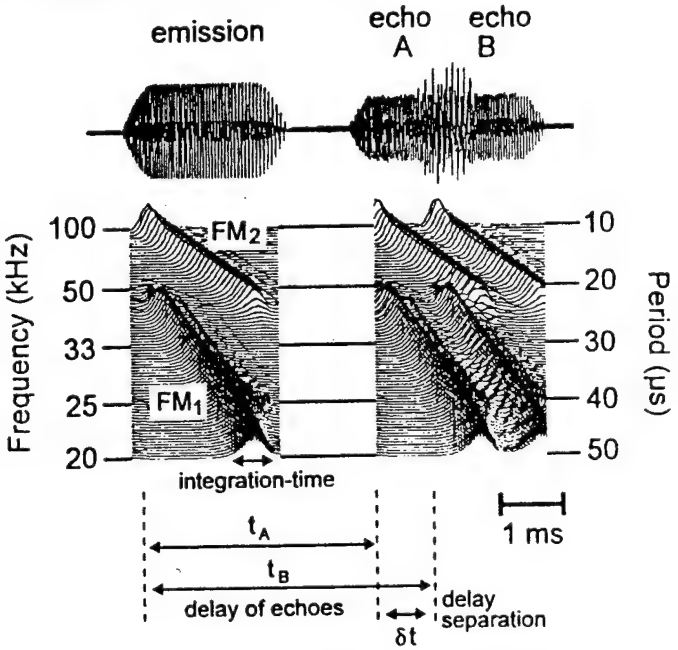

Fig. 5. (Top) Waveforms of 2-ms two-harmonic FM sonar emission and two echoes (A,B) at delays of 3.7 and 4.7 ms. (Bottom) SCAT spectrograms representing hyperbolic frequency axis for emission and two echoes at top (81 parallel band-pass-filter channels followed by rectification and smoothing with integration-time of 300-400 μs). Although the raw echo waveforms overlap, their spectrograms are separate because the delay separation (1 ms) is more than the integration-time (300-400 μs).

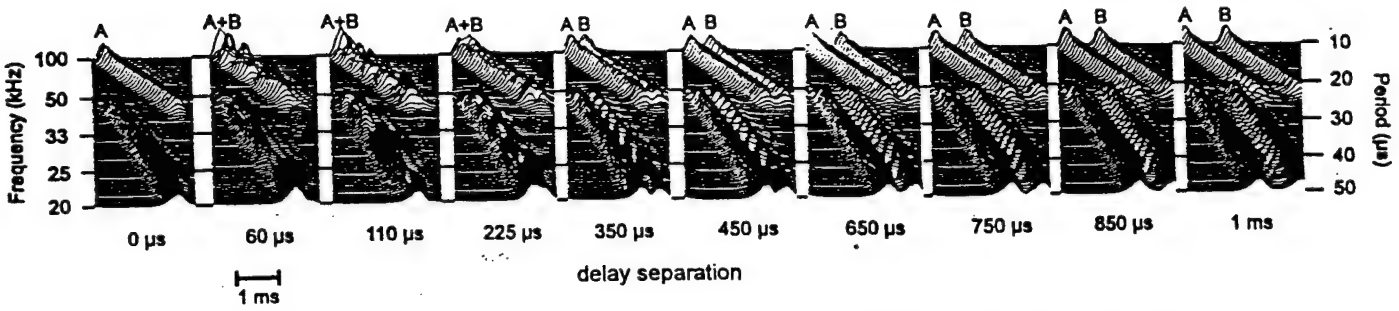

Fig. 6. SCAT (hyperbolic-frequency) spectrograms of pairs of overlapping echoes (A,B from Fig. 5) at different delay separations in relation to their integration-time of 300-400 µs. They appear as separate spectrograms for separations larger than the integration-time, and merge into the same spectrogram for separations that are smaller.

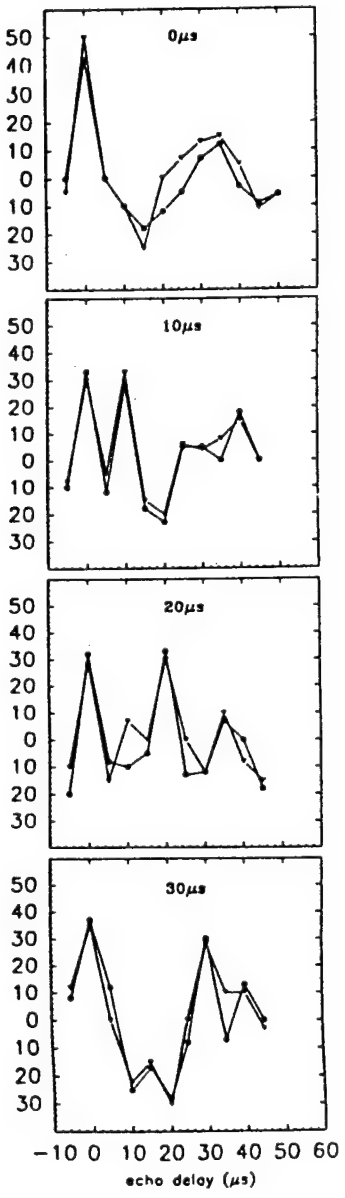

Fig. 7. Graphs showing the compound performance curves (% errors) for two *Eptesicus* in a task that uses a probe echo at different delays (horizontal axis) to locate the delay values perceived for two overlapping test echoes (similar to A and B in Fig. 6) at delay separations of 0, 10, 20, or 30 µs. Bats make errors (peaks) at delays where they perceive test and probe echoes to have same delay (60 trials per data-point per bat). Each test echo is perceived at its correct delay--the target has two glints, and the image has two glint components (0 µs on horizontal axis is equivalent to 3.2-ms overall delay of nearer of the two overlapping echoes).

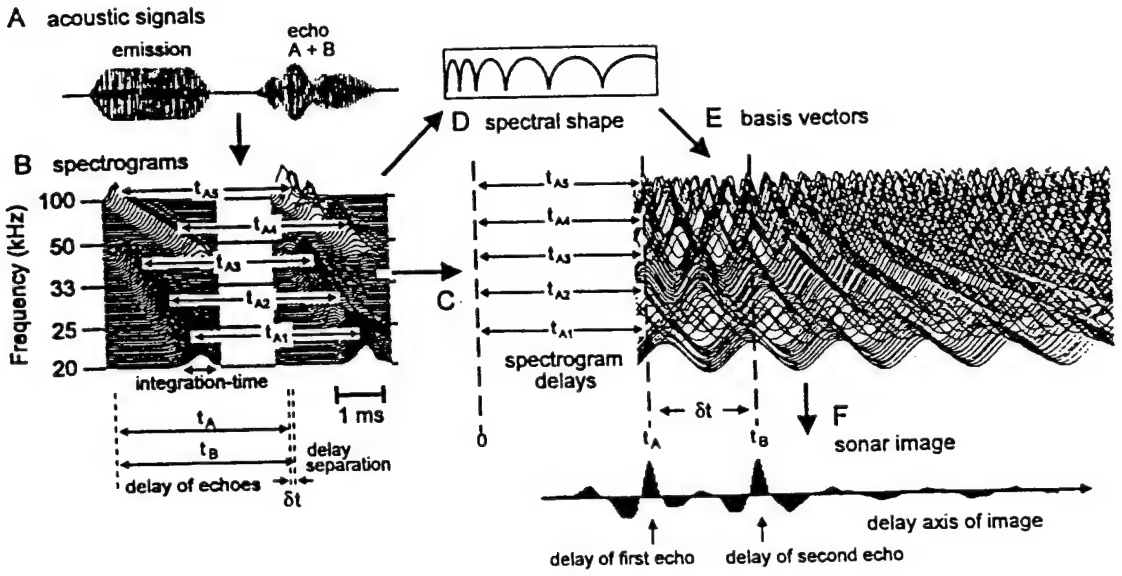

Fig. 8. Diagram of SCAT computational model of convolution-deconvolution echo-processing for target range and fine range separation by *Eptesicus* (Saillant et al. 1993). (A) waveform of FM sonar emission and two echoes (A+B) separated by 60 μs. (B) 81-channel SCAT spectrograms of waveforms in A, with spectrogram delays (t_{A1}-t_{A5}) delineating time-offset of echo. (C) Alignment and averaging of spectrogram delays to determine overall delay (indicated as t_A). (D) Shape of spectrum for overlapping echoes, with notches and peaks caused by interference (see Section 4.4.1). (E) Cosine-phase basis vectors individually phase-aligned to start at time specified by spectrogram delay for each channel. (F) Echo-delay ("A-Scope") sonar image formed by summing basis vectors. Image contains delay for echo A originally from spectrogram delays and echo B from summation and cancellation of different basis vectors phase-aligned to each channel of the spectrogram. The shapes of the image components resemble emission-echo crosscorrelation functions weighted by the hyperbolic frequency regime.

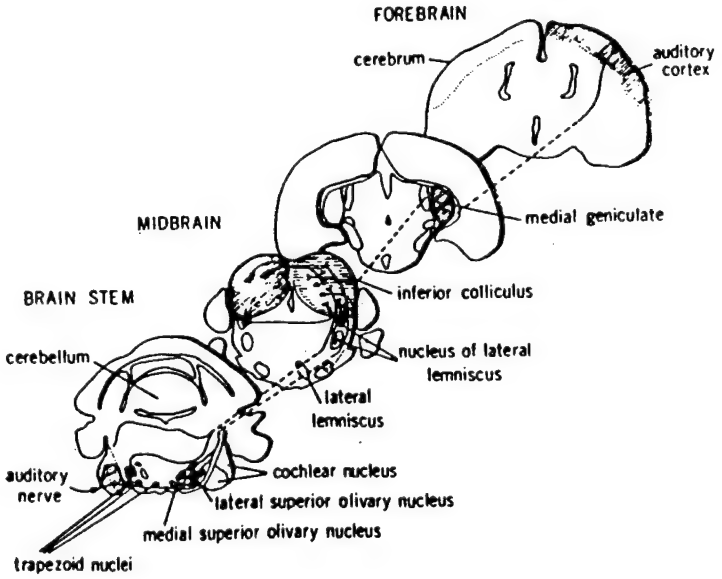

Fig. 9. Diagram showing principal auditory centers in the bat's brain. These anatomical structures receive their inputs from auditory stimulation approximately in succession (after Schweizer 1981).

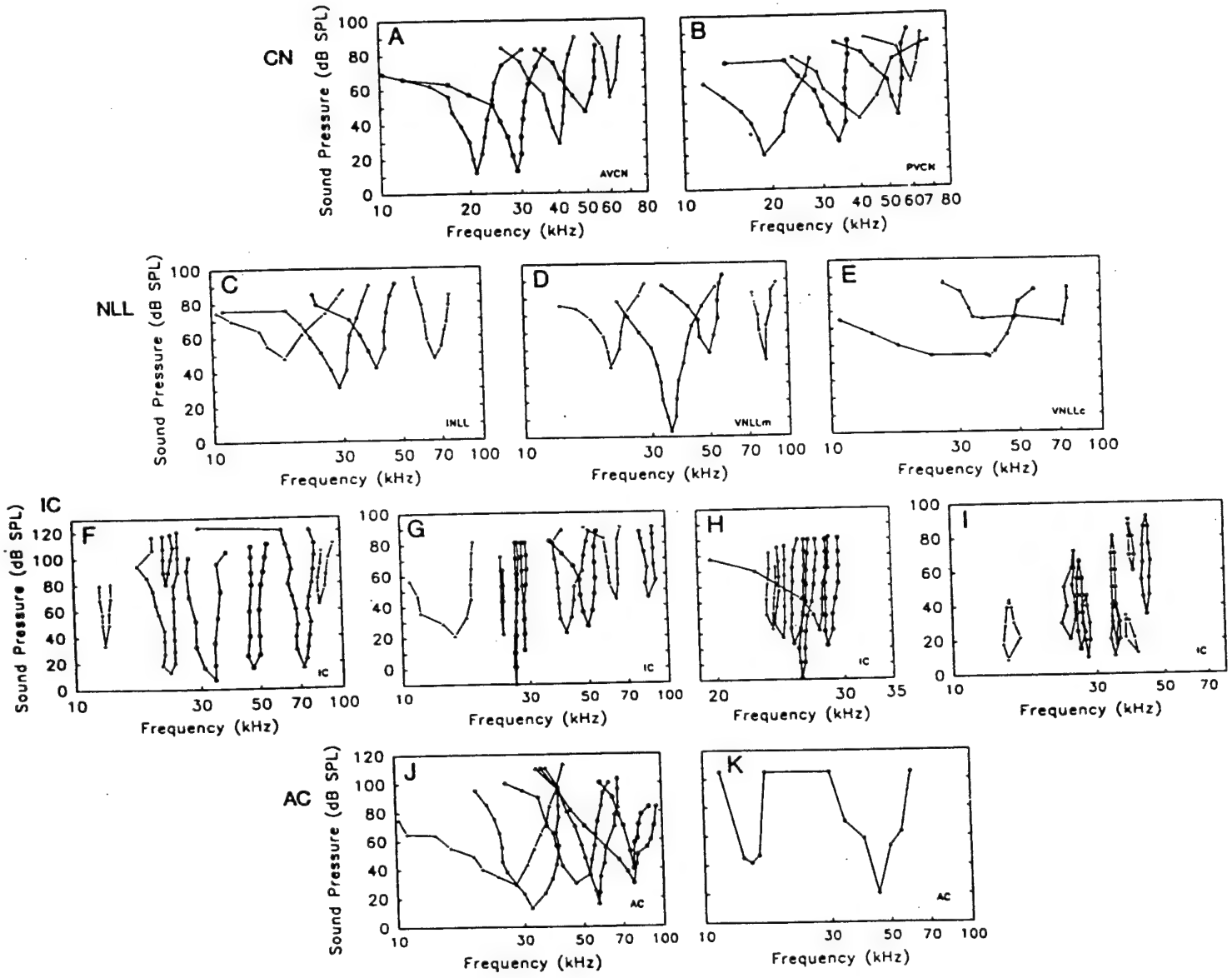

Fig. 10. Representative frequency tuning curves for *Eptesicus* from (A-B) the cochlear nucleus (CN), (C-E) the nucleus of the lateral lemniscus, (NLL), (F-I) the inferior colliculus (IC), and (J-K) the auditory cortex (AC). (see Fig. 9.)

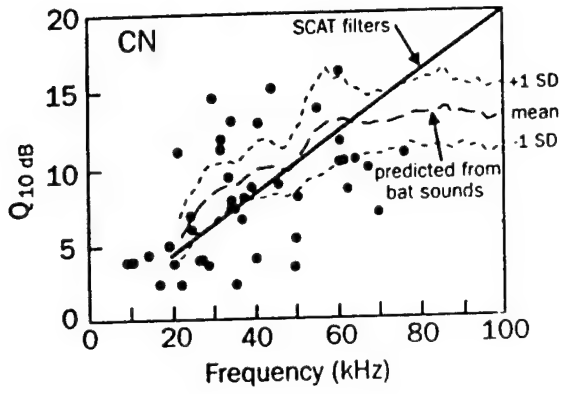

Fig. 11. Sharpness of frequency tuning (Q_{10dB}) in the cochlear nucleus of *Eptesicus*. (Data-points are from Haplea, Covey, and Casseday 1994; dashed lines are predicted tuning values--mean ±1 SD--from Menne 1988; solid sloping line is tuning of SCAT filters measured from frequency-response curves--Saillant et al. 1993.)

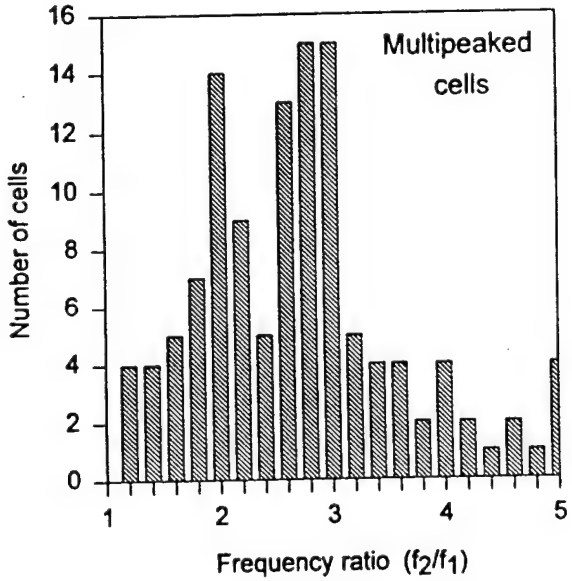

Fig. 12. Distribution of ratios for higher and lower frequencies (f_2/f_1) in auditory cortical neurons with two widely-separated tuned frequencies (f_1,f_2). Ratios are clustered around 2:1 and 3:1.

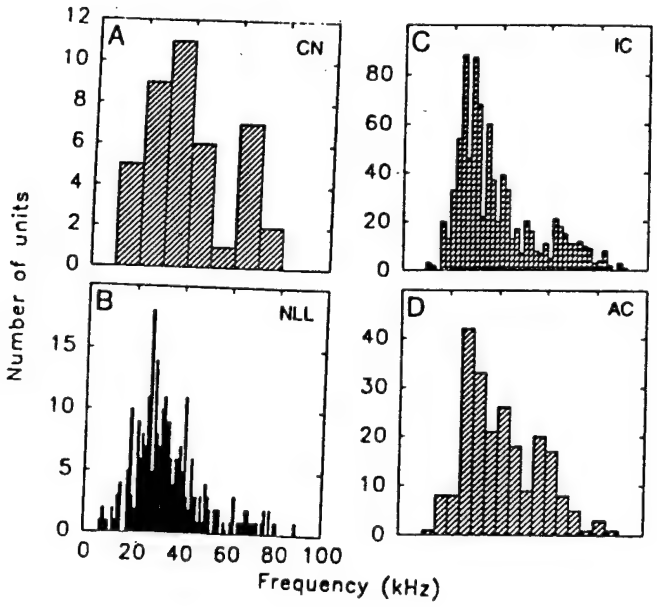

Fig. 13. Density of frequency tuning in *Eptesicus* at different frequencies for the cochlear nucleus (CN), the nucleus of the lateral lemniscus (NLL), the inferior colliculus (IC), and the auditory cortex (AC). Tuning is predominantly at frequencies of 25-50 kHz and secondarily at 60-70 kHz.

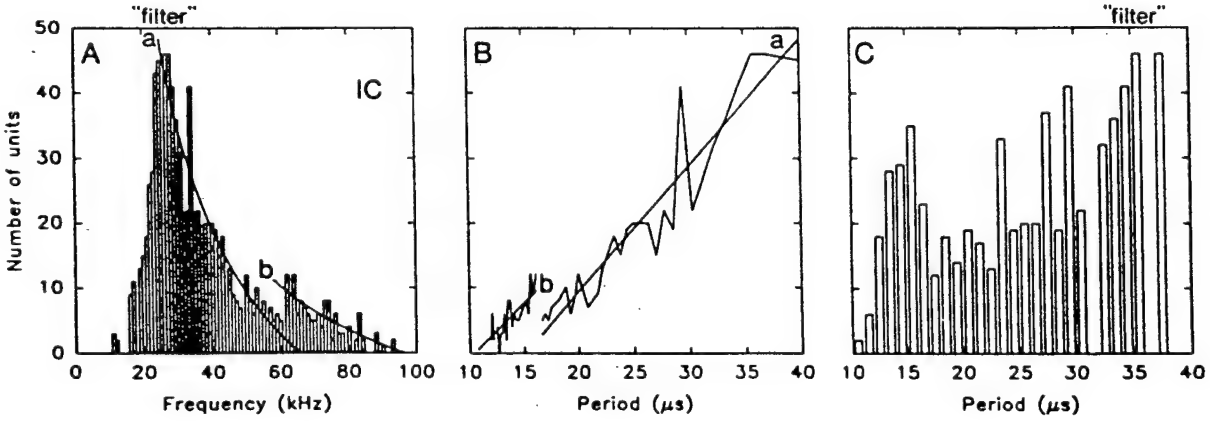

Fig. 14. Detailed plots showing density of frequency tuning in the inferior colliculus. (A) Density of frequency tuning at different frequencies. (B) Density of frequency tuning at different periods (regression lines a,b and curves a,b from this graph). (C) Density of *period* tuning. The density profile is more nearly uniform across different periods (C) than across different frequencies (A), indicating that the sampling domain for frequency is roughly hyperbolic ("filter" region discussed in text).

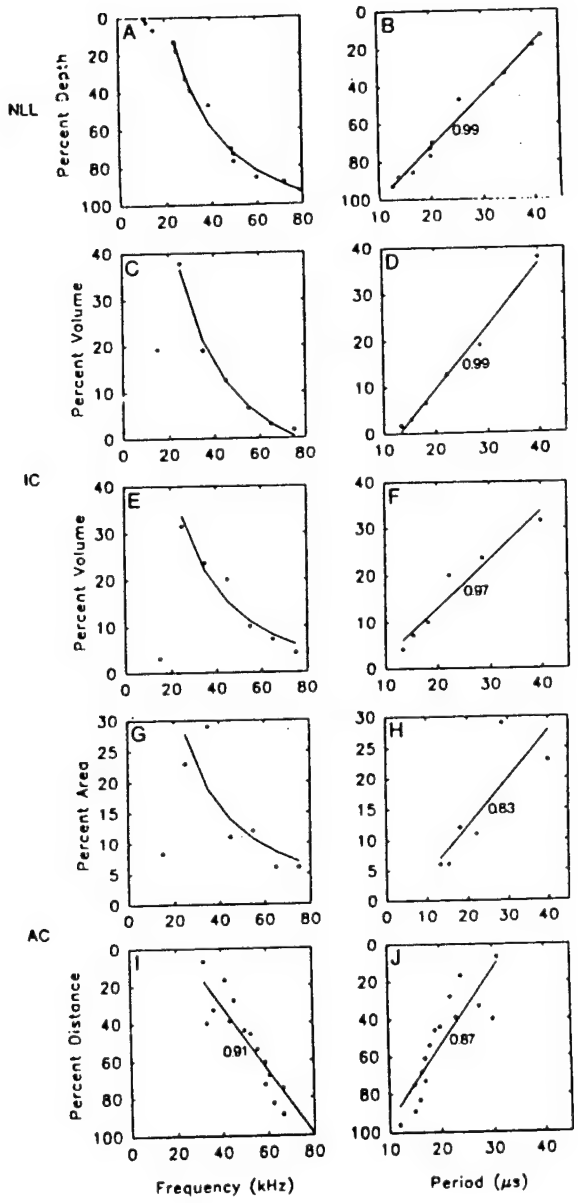

Fig. 15. Graphs showing aspects of tonotopic organization in the nucleus of the lateral lemniscus (NLL), the inferior colliculus (IC), and the auditory cortex (AC). A,C,E,G,I show anatomical location plotted against frequency, B,D,F,H,J show anatomical location plotted against period.

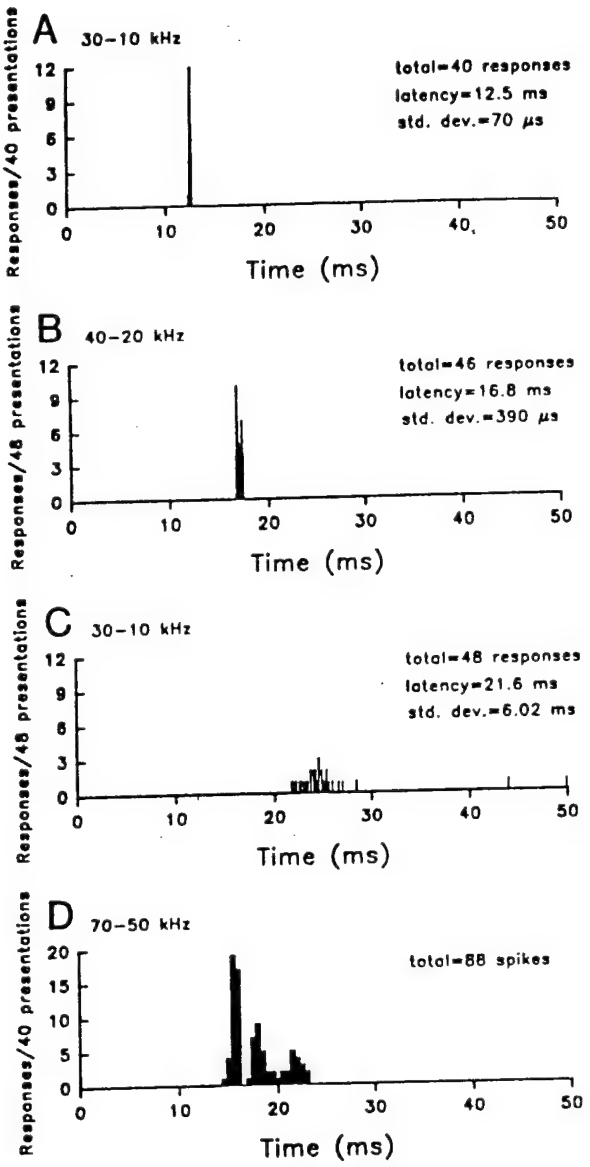

Fig. 16. Responses of neurons in the inferior colliculus of *Eptesicus* to 2-ms FM sounds. (A,B,C) Latency (or PST) histograms of on-responses in three different neurons evoked by 40 presentations of a 2-ms FM sweep that covered each cell's tuned frequency. 93% of these neurons discharge only once per short-duration FM stimulus, but they differ in their characteristic latency and latency variability (standard deviation). (D) Latency histogram for multi-unit response dominated by on-discharges in three neurons with latencies at a periodic spacing.

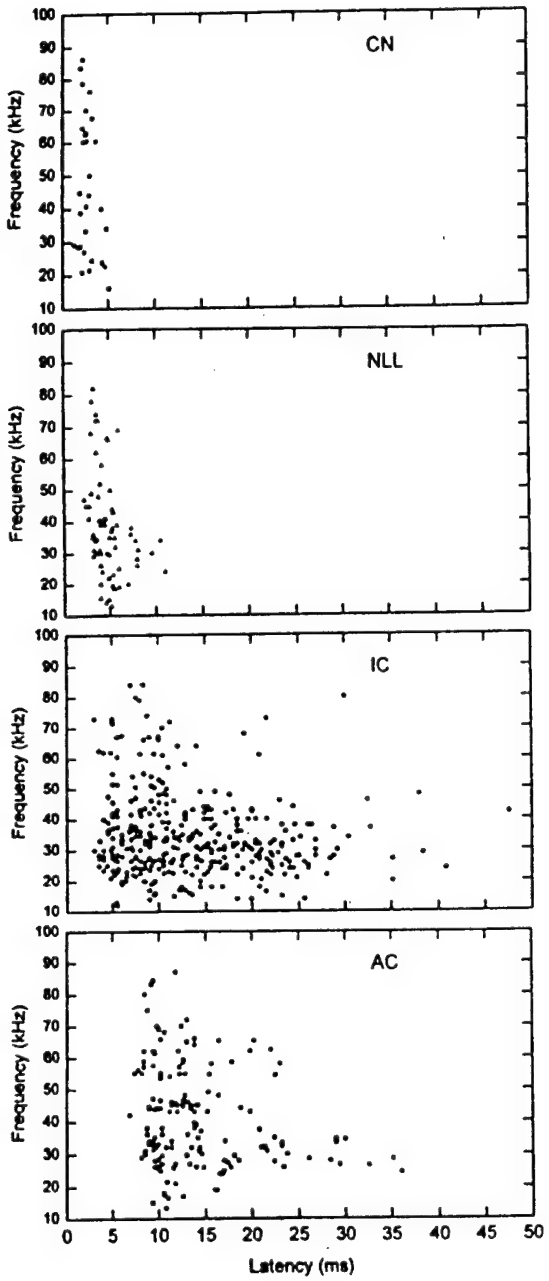

Fig. 17. Latencies of on-responses at different tuned frequencies in the cochlear nucleus (CN), the nucleus of the lateral lemniscus (NLL), the inferior colliculus (IC), and the auditory cortex (AC). Responses in lower auditory centers (CN, NLL) have a narrow spread of latencies, but at higher centers (IC, AC) they are dispersed up to 30-35 ms.

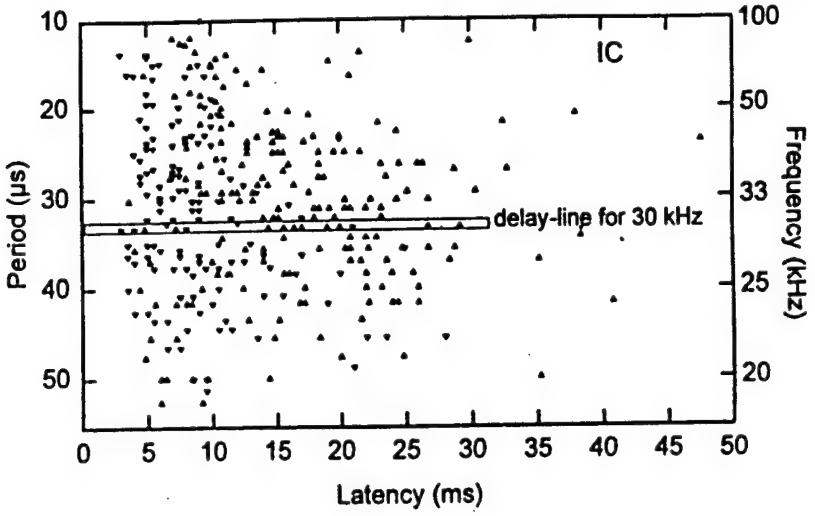

Fig. 18. Latencies of on-responses in the inferior colliculus replotted on a hyperbolic frequency axis to correspond to SCAT spectrograms (Fig. 5).

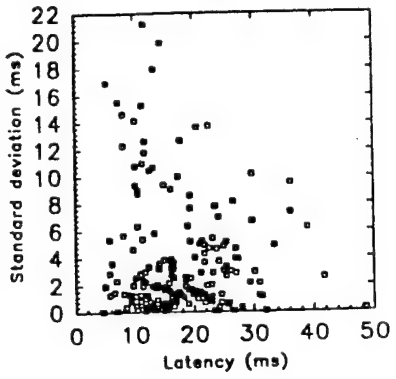

Fig. 19. Graph showing the distribution of latency variability (standard deviation) for on-responses in the inferior colliculus having different characteristic latencies. There are cells with latencies from 4 ms to 25-30 ms that have narrow variability (standard deviations of 50-100 µs or less).

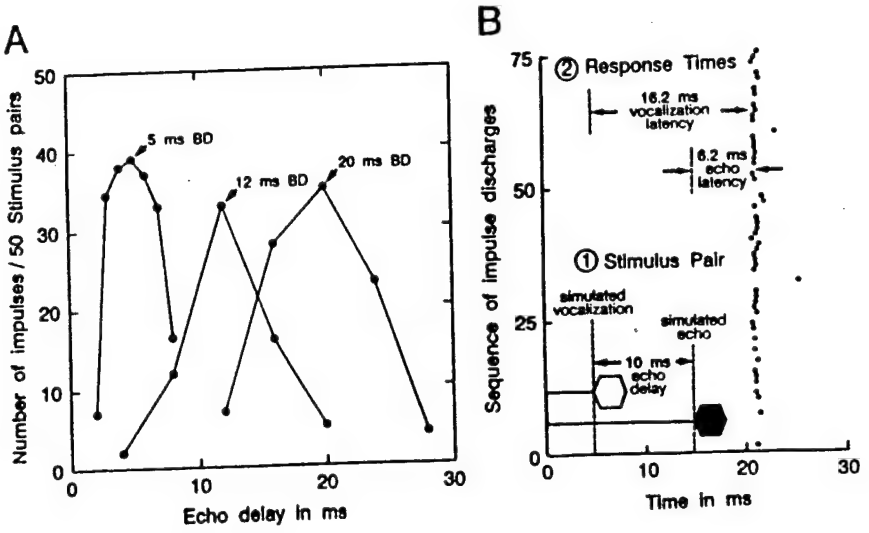

Fig. 20. Delay tuning in the auditory system of *Eptesicus*. (A) Delay-tuning curves for three different cortical neurons with best delays of 5, 12, and 20 ms. Even though the bat's echo-delay acuity is a fraction of a microsecond, the width of delay-tuning curves is in the range of milliseconds. (B) Dot-raster plot of 75 successive delay-tuned responses to an echo at the 10-ms best delay of a neuron, showing the latency from the emission, the latency from the echo, and the difference, which corresponds to best delay. The variability of response latency is only 300 μs, while this cell's delay-tuning curve is 7.5 ms wide. Even after delay-tuning is established, the timing of responses still conveys needed information about delay.

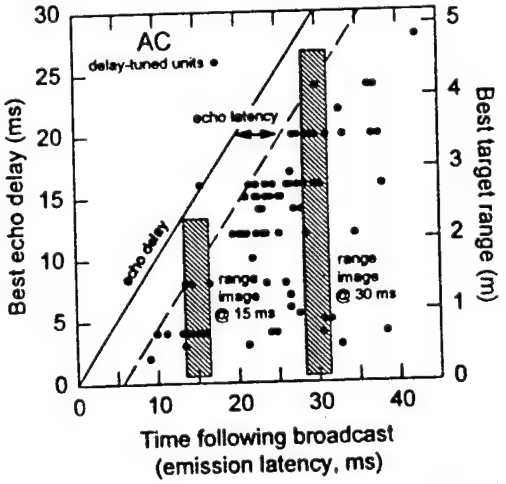

Fig. 21. Graph showing the emission latency of cortical delay-tuning at different values of best delay. The echo arrives at delays indicated by upward-sloping solid line, and cortical responses begin to occur as early as 6 ms afterward along sloping dashed line. As time progresses, echoes return from targets at longer ranges, to be registered by delay-tuned responses, but delay-truned responses to targets at shorter ranges that are already registered continue to occur, storing older information about near targets while newer information builds up about far targets. Vertical shaded bars delineate best delays in range images occurring 15 and 30 ms after the broadcast.

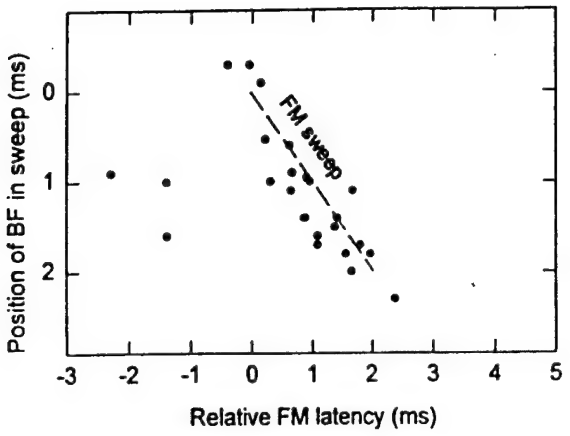

Fig. 22. Relative latencies of on-responses in inferior colliculus neurons tuned to different frequencies reconstruct the location of that tuned frequency along FM sweeps. Latencies of on-responses to tone-bursts at each cell's tuned frequency are subtracted from FM latencies and then corrected for the periodic spacing of responses to tone-bursts relative to FM sweeps. The spectrogram of the sweep is concealed within the dispersed absolute latencies of the responses (Fig. 18). (see also Bodenhamer and Pollak 1981.)

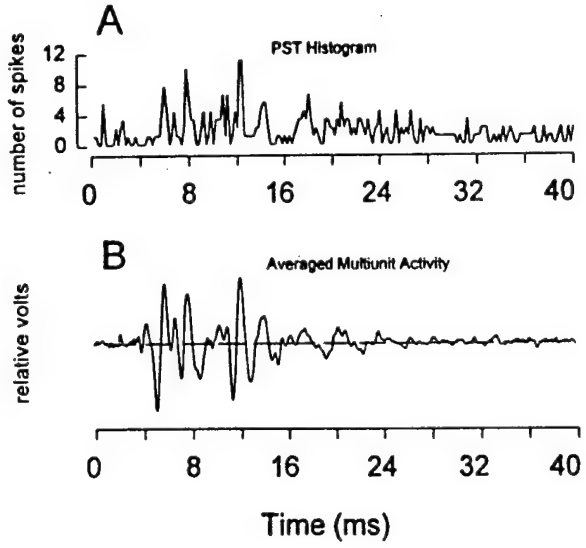

Fig. 23. Comparison of latency histogram (A) and analog-averaged response (B) for multi-unit response from the inferior colliculus of *Eptesicus*. The histogram (vertical scale is number of discharges in each 50-μs time-bin for 40 stimulus repetitions) and the average (vertical scale is voltage, with maximum peak-height about 100 μV at electrode tip; N=256) both show the same prominent features, but the relative noisiness of the histogram (see Section 7.2.3) conceals finer details and even partially obscures the larger peaks. For responses originating from local clusters of units, the averaged response is a superior index of latency organization, which is composed of on-responses in different cells with periodically-staggered latencies.

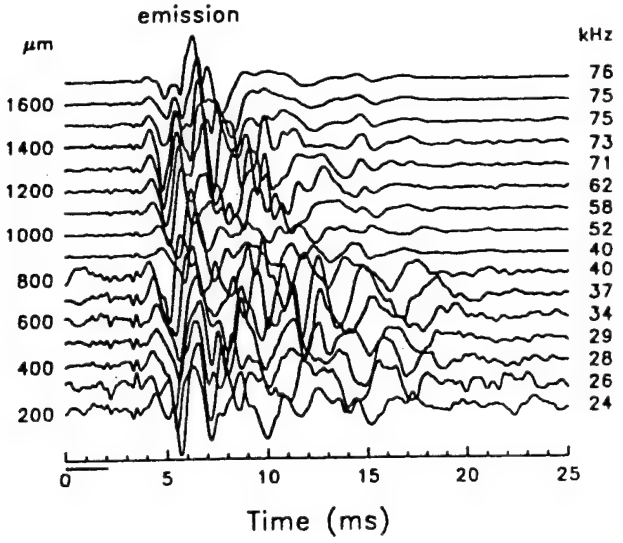

Fig. 24. Series of local averaged multi-unit responses (N=256) recorded in the inferior colliculus from different depths (left vertical axis identifying series of responses) tuned to different frequencies (right vertical axis). The stimulus is a 2-ms multiple-harmonic FM sound simulating a sonar emission of *Eptesicus* (horizontal bar below origin of time axis). Duration of the responses and range of latencies for response-peaks correspond to dispersed latencies of single-unit on-responses (Fig. 18). The responses exhibit organization across tuned frequencies consisting of periodically-spaced ridges that slope downward to the right, with slopes that change gradually as latency increases.

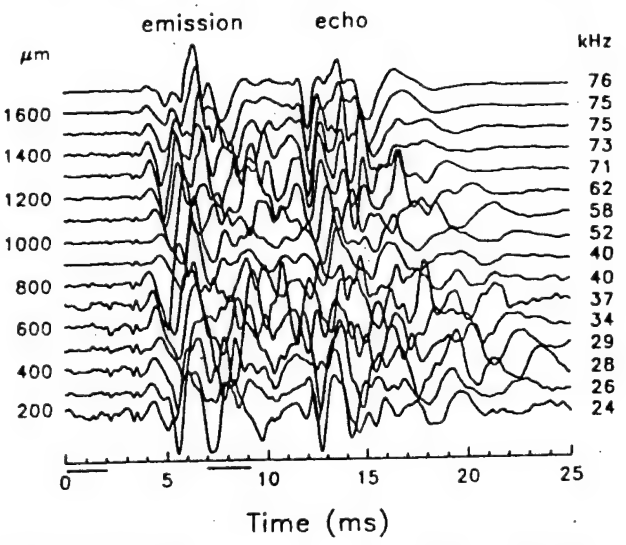

Fig. 25. Series of local averaged multi-unit responses (N=256) to a pair of 2-ms multiple-harmonic FM sounds simulating a sonar emission of *Eptesicus* (horizontal bar below origin of time axis) and an echo at a delay of 7 ms (horizontal bar below 7-ms point on time axis). Both sounds evoke pattern of responses similar to that shown on Fig. 24, with overlap of their duplicate series of ridges at emission latencies from about 12 ms to about 18-20 ms. These responses have their own physiological time-frequency space (horizontal and vertical axes identifying responses) related to time-of-occurrence and frequency of FM sounds; they constitute a type of signal representation that displays echo delay in the overlap of the response patterns.

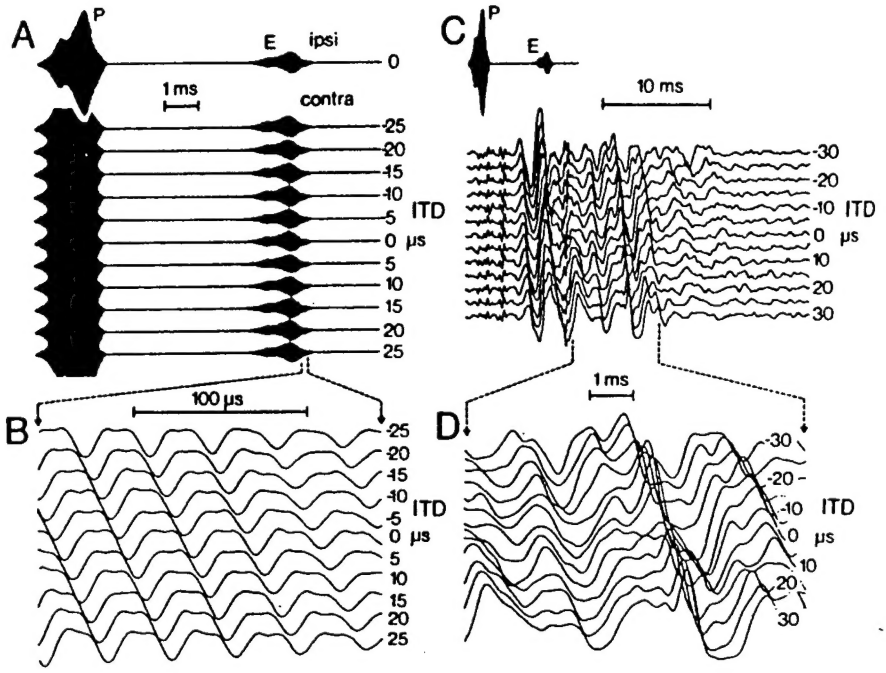

Fig. 26. The latency structure of local averaged multi-unit responses (N=256) in the inferior colliculus encodes information about fine temporal structure of echoes, in this example binaural delay or phase differences. (A) Envelopes of stimuli that mimic 2-ms FM emission and echo (6-ms overall echo delay, with changes from -25 µs to +25 µs in delay of contralateral echo to mimic changes in target azimuth). (B) Expanded view of contralateral echo waveform to show shifts in time or phase in one ear relative to the other. (C) Averaged responses recorded from site tuned to 20-22 kHz for series of binaural echo-delay differences from -30 µs to +30 µs (stimulus envelopes on same time-scale at top). Latencies of responses shift to right by about 100 times the binaural delay difference in the echo. (D) Expanded view of multi-unit responses to show detailed structure and how it resembles a time-expanded reconstruction of echo acoustic waveform in B. (Note different time scales for each part of figure.)

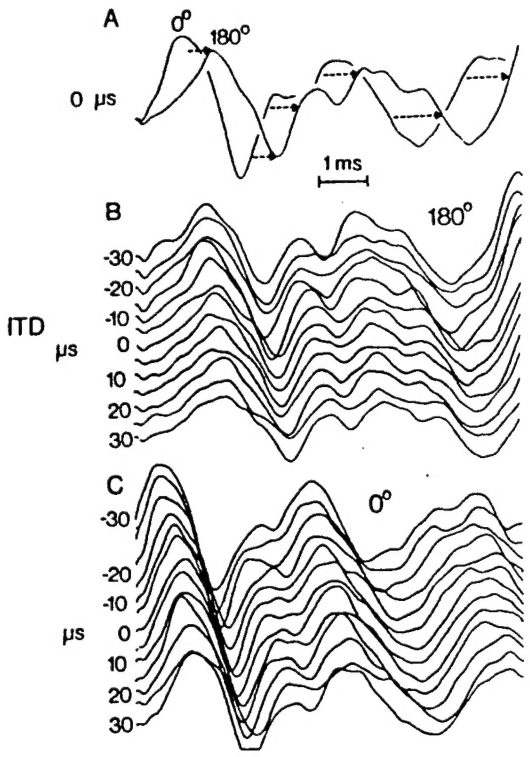

Fig. 27. The latency structure of local multi-unit responses (N=256) in the inferior colliculus encodes information about echo phase. (A) Segments of responses (same as D in Fig. 26) showing latency shifts of peaks for 180° echo phase-shift. Binaural echo-delay difference is 0 µs. (B,C) Series of responses for echoes at 0° or 180° phase at different binaural delay differences from -30 µs to +30 µs to show joint coding of binaural delay difference and phase on expanded time scales.

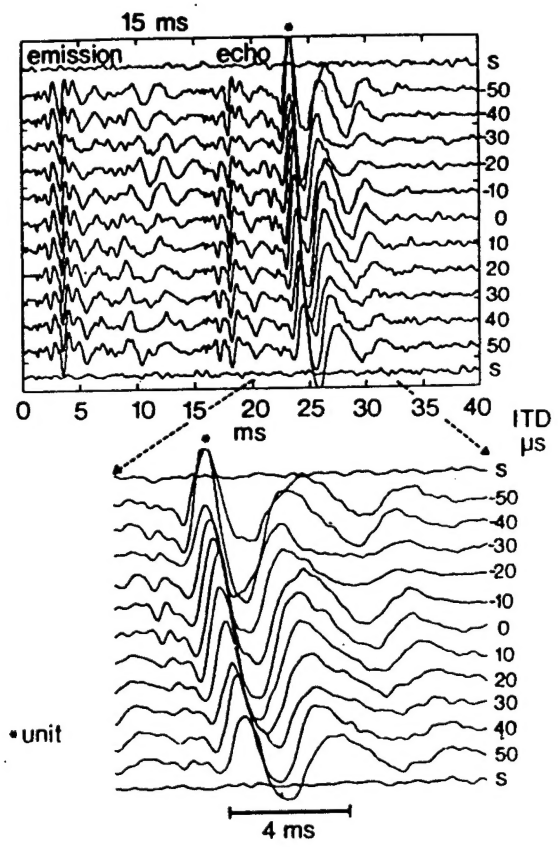

Fig. 28. The latency structure of local multi-unit responses (N=256) in the auditory cortex encodes overall echo delay on a coarse scale and binaural delay or phase on an expanded time scale. (Top) Responses averaged from recording of a delay-tuned cortical neuron to a 2-ms FM emission followed by an echo at the cell's best delay (15 ms) at different binaural delay differences from -50 μs to +50 μs. Peak marked * at emission latency of about 23-24 ms is cortical response; earlier peaks are brain-stem and midbrain responses to the emission (latency of 2-5 ms) and echo (latency of 16-18 ms) that "leak" into recording from a distance. Traces labeled S show spontaneous activity in absence of stimulus. (Bottom) Expanded view of cortical response to show size of latency shift induced by small binaural delay differences. This cell's delay-tuning curve determined from its isolated discharges as a single unit represents delay on a scale of milliseconds (see Fig. 20), while the latency of the local averaged response, which contains contributions from neighboring cells as well, encodes binaural delay differences on an expanded time scale of microseconds. These responses also undergo latency shifts in response to 0° or 180° echo phase-shifts (similar to effect in Fig. 27).

Fig. 29. The latency structure of local multi-unit responses (N=128) in the auditory cortex encodes information about fine delay separation of overlapping echoes within the bat's 350-µs integration-time for echo reception. Graph shows a series of responses over a 25-ms span of emission latencies (horizontal axis shows time after the onset of the envelope of the 12-ms FM emission and also the very beginning of the echo at a delay of 24 ms) for echoes simulating two reflected replicas at delay separations of 0-200 µs at 25-µs steps. Main peaks in responses shift in latency by about 3 ms for the total 200-µs span of delay separations, with a larger initial latency shift of about 3.5 ms from delay-separation of 0 µs to 25 µs. This series of responses demonstrates an expanded time-domain representation of information that initially was represented by the spectrum of the overlapping echoes.